1-1-2016

Essays on Labor and Risk

Patrick E. Dejarnette
University of Pennsylvania, pdejar@wharton.upenn.edu

Essays on Labor and Risk

Abstract
This dissertation presents three essays in labor economics and risk. Chapter 1 examines how past effort can impact current effort, such as when effort is reduced following an interruption. I present a series of real-effort incentivized experiments in which both piece rates and leisure options were manipulated and find effort displays significant stickiness, even in the absence of switching costs. I demonstrate that this intertemporal evidence is indicative of effort "momentum", rather than on-the-job learning, reciprocity, or income targeting. When employing an instrumental variables (IV) approach, approximately 50\% of the effort increase persists for 5 minutes after incentives return to baseline. Thus if a worker suffers a complete interruption in productivity, it would take an average of 15 minutes to return to 90\% of prior work effort. I further demonstrate that advanced knowledge does not significantly reduce this productivity loss.

Chapter 2 examines how risk preferences differ over goods and in-kind monetary rewards. I study an incentivized experiment in which subjects allocate bundles of either Amazon.com goods or Amazon.com gift credit (which must be spent immediately) across uncertain states. Under a standard model of perfect information of prices and goods available, I demonstrate risk preferences across these treatments would be identical. In practice, I uncover substantial differences in risk preferences across goods and in-kind monetary rewards. With additional treatments, I find no evidence that these differences are driven by price or product uncertainty.

Chapter 3 is joint work with David Dillenberger, Daniel Gottlieb, and Pietro Ortoleva. We study preferences over lotteries that pay a specific prize at uncertain dates. Expected Utility with convex discounting implies that individuals prefer receiving x in a random date with mean t over receiving x in t days for sure. Our experiment rejects this prediction. It suggests a link between preferences for payments at certain dates and standard risk aversion. Epstein-Zin (1989) preferences accommodate such behavior, and fit the data better than a model with probability weighting.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Applied Economics

First Advisor
Jeremy Tobacman

Keywords
effort momentum, risk over goods, time lotteries

Subject Categories
Economics | Economic Theory | Labor Economics

ESSAYS ON LABOR AND RISK

Patrick E. DeJarnette

A DISSERTATION

in

Applied Economics

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

____

Jeremy Tobacman
Assistant Professor of Business Economics and Public Policy

Graduate Group Chairperson

____

Eric Bradlow, K.P. Chao Professor; Professor of Marketing, Statistics, and Education

Dissertation Committee:

Iwan Barankay, Associate Professor of Business Economics and Public Policy
Daniel Gottlieb, Assistant Professor of Economics, Washington University in St. Louis
Robert Jensen, Professor of Business Economics and Public Policy
Judd Kessler, Assistant Professor of Business Economics and Public Policy

To my family,

to whom I owe everything.

ACKNOWLEDGEMENT

This dissertation would not have been possible without the support, guidance, and encouragement of many people. First and foremost, my tireless dissertation committee, Jeremy Tobacman, Iwan Barankay, Daniel Gottlieb, Robert Jensen, and Judd Kessler. Jeremy's persistence and patience, Iwan's advice, Daniel's encouragement, Rob's guiding hand, and Judd's insight have all played a crucial role in getting me to where I am today. I am especially thankful to my primary advisor Jeremy who has met with me weekly for over 4 years and has weathered both the ups and downs of the research process. Without his extraordinary efforts, I would not have weathered this storm.

I am indebted to the many readers who gave me feedback on my dissertation at various stages. Including comments from Paul Bruton, Mark Clements, Amanda Chuan, Ulrich Doraszelski, Clayton Featherstone, Fernando Ferreira, Joseph Harrington, Sonia Jaffe, Corinne Low, Anita Mukherjee, Todd Sinai, Charles Sprenger, Shing-Yi Wang, Kate Wilhelm, and Tomoyoshi Yabu. I am also grateful to my excellent co-authors, David Dillenberger, Daniel Gottlieb, and Pietro Ortoleva, without whom Chapter 3 would not have been possible. This research (specifically Chapter 1 and 2) was also supported by a grant from the Wharton School's Mack Institute for Innovation Management.

More than anyone though, Tian Cai had the largest impact on my experience in graduate school. To her I owe my deepest gratitude for supporting me and serving as my guiding star. Her incredible love and remarkable insight can be found on every page of this dissertation. I thank her for inspiring me and for always being my better half.

Finally, I owe everything to the rest of my family, particularly my mother Mary Pat, my father Edward, my brother Robert, and my sister Stephanie. Without all of them, I would never have reached where I am today. Their support and insistence have made me the person I am today, for which I am eternally thankful and proud.

ABSTRACT

ESSAYS ON LABOR AND RISK

Patrick E. DeJarnette

Jeremy Tobacman

This dissertation presents three essays in labor economics and risk. Chapter 1 examines how past effort can impact current effort, such as when effort is reduced following an interruption. I present a series of real-effort incentivized experiments in which both piece rates and leisure options were manipulated and find effort displays significant stickiness, even in the absence of switching costs. I demonstrate that this intertemporal evidence is indicative of effort momentum, rather than on-the-job learning, reciprocity, or income targeting. When employing an instrumental variables (IV) approach, approximately 50% of the effort increase persists for 5 minutes after incentives return to baseline. Thus if a worker suffers a complete interruption in productivity, it would take an average of 15 minutes to return to 90% of prior work effort. I further demonstrate that advanced knowledge does not significantly reduce this productivity loss.

Chapter 2 examines how risk preferences differ over goods and in-kind monetary rewards. I study an incentivized experiment in which subjects allocate bundles of either Amazon.com goods or Amazon.com gift credit (which must be spent immediately) across uncertain states. Under a standard model of perfect information of prices and goods available, I demonstrate risk preferences across these treatments would be identical. In practice, I uncover substantial differences in risk preferences across goods and in-kind monetary rewards. With additional treatments, I find no evidence that these differences are driven by price or product uncertainty.

Chapter 3 is joint work with David Dillenberger, Daniel Gottlieb, and Pietro Ortoleva.

We study preferences over lotteries that pay a specific prize at uncertain dates. Expected Utility with convex discounting implies that individuals prefer receiving x in a random date with mean t over receiving x in t days for sure. Our experiment rejects this prediction. It suggests a link between preferences for payments at certain dates and standard risk aversion. Epstein-Zin (1989) preferences accommodate such behavior, and fit the data better than a model with probability weighting.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Effort Momentum

1.1. Introduction

By some estimates, interruptions disrupt 1.5 to 2.1 hours per work day for over 56 million US "knowledge workers" (Gonzalez and Mark (2004); Spira and Feintuch (2005)). Observational studies show that hospital workers are interrupted 5 times per hour (Weigl et al. (2014); Berg et al. (2013)) while software developers and managers are interrupted 25 times per day (Gonzalez and Mark (2004)).[1] In similar studies, 15-23% of interrupted work is not resumed on the same day, a particular concern within the health services literature (Westbrook et al. (2010); Mark et al. (2005)). However, this evidence is difficult to interpret when the interruptions themselves may be necessary, as within a hospital's emergency department. In other contexts, interruptions from a manager may reduce principal-agent concerns through increased monitoring or communication as formulated by Coviello et al. (2014). Interruptions from a co-worker could increase a firm's total output, even at the expense of the interrupted worker. As a result, it might be presumptuous to target interruptions themselves as a source of productivity waste.

At the heart of the question of how interruptions affect behavior is whether there is stickiness in effort allocation. If so, interruptions could undermine productivity due to unplanned effort reduction following an interruption. This paper answers this more general question of whether productivity loss persists over time, and if so, what might be done to recover it.

This "loss of momentum" is often posited by the psychology literature,[2] media,[3] and consulting reports,[4] but has not been thoroughly examined within the economics literature.

[1] While half of these are considered "internal" interruptions, which may be more indicative of task juggling (Coviello et al. (2014)), the remaining half of "external" interruptions are arguably the most time intensive. For example, the data in Gonzalez and Mark (2004) also shows that 1.5 hours per day are spent on unscheduled meetings such as workers stopping by or talking through cubicle walls.

[2] See an extensive psychology literature on "flow", cf. Nakamura and Csikszentmihalyi (2002); Schaffer (2013), in which clear tasks with adequate challenge and objective goals enables a continuous work state.

[3] E.g. "*First, there is the diversion itself, taking your employees off task after they have assembled the resources and thinking necessary for that particular task. Then there is the restart—reassembling the resources, thoughts, and readiness. There is the loss of momentum caused by the initial distraction from the original purpose.*" (Brown (2015))

[4] Such as " *.. even a one-minute interruption can easily cost a knowledge worker 10 to 15 minutes of*

The closest literature on intertemporal labor supply tends to focus on longer time scales and hours worked rather than output (Camerer et al. (1997); Oettinger (1999); Farber (2005, 2008); Crawford and Meng (2011); Chetty et al. (2011)). Within this literature, Fehr and Goette (2007) employ a field experiment on bicycle riders and finds evidence consistent with a model in which past effort exhausts riders, making additional effort more costly. Under this model, an exogenous interruption in effort could actually boost future productivity as the worker has had a chance to "catch their breath".[5] However, the longer time scale and the physical nature of the task make these findings difficult to apply to interruptions among a broader class of knowledge workers.[6]

In this paper, I hypothesize and test a theory in which past effort has a direct effect on disutility from present effort. This model has theoretical similarities to a model of habit preferences, but over effort rather than consumption. I refer to this theory as effort momentum.

To test for the presence of effort momentum, I conduct a series of real-effort laboratory experiments with 577 University of Pennsylvania students at the Wharton Behavioral Lab. This controlled setting allows me to observe workers' responses to both piece rates and leisure opportunities over multiple periods. The workers complete counting or slider tasks on a computer screen but have the option to engage in leisure by viewing YouTube videos at any time.[7] I manipulate (i) the piece rate for completed problems and (ii) the leisure opportunities available (by varying subjects' access to their cell phones). Subjects are quizzed prior to every period to ensure incentives and leisure options are understood.

---

lost productivity due to the time needed to reestablish mental context and reenter the flow state." (Nielsen (2003))

[5]While consistent with an effort exhaustion model, the authors find that individual measures of loss aversion are predictive of the effort decrease, suggesting a model of loss-aversion may be more appropriate for that setting.

[6]There is a noted lack of consensus regarding what constitutes a "knowledge worker", but in line with the literature, I characterize knowledge work "as less tangible than manual work and using the worker's brain as the means of production" (Ramírez and Nembhard (2004)). This would roughly coincide with the 56 million "management, professional, and related occupations" from the 2014 Current Population Survey.

[7]This is in line with recent work demonstrating the importance of outside leisure options for external validity of laboratory experiments. (Corgnet et al. (2014); Charness et al. (2010); Eriksson et al. (2009); Kessler and Norton (2015))

The laboratory setting for the experiment enabled me to accurately measure productivity. This accuracy allows me to induce variation in effort over short time scales by changing incentives quickly. In addition, while the tasks involved are somewhat artificial, exerting effort on a computer located in a cubicle closely resembles a relevant work environment for many knowledge workers (Gonzalez and Mark (2004)).[8] The laboratory also eliminates peer effect confounds that might be present in a field setting, such as fairness concerns over some workers being paid more. Moreover, the setting allowed me to replicate across multiple designs and tasks to ensure effort momentum is not limited to a single context. The laboratory environment also made it possible to accurately enforce available leisure opportunities, particularly cell phone access. This leisure variation is important for differentiating momentum from alternate theories such as reciprocity.

My results show significant evidence of effort stickiness. Workers treated with a higher (lower) piece rate exert more (less) effort in the treated period relative to control.[9] Even after financial incentives return to baseline, workers who received a higher piece rate continue to work harder than those who only received a baseline piece rate. This lingering effort differential is approximately half of the original effort increase induced by the heightened piece rate. By the same token, workers who receive a lower piece rate in one period continue to exert less effort in following periods relative to the control group. These findings indicate that effort allocation in one period may depend positively on recent work effort.

This evidence of effort stickiness could be a result of momentum, reciprocity, on-the-job learning, or potentially other interpretations. To identify the source of this effort stickiness, I structured the experimental design to provide additional comparisons informed by theoretical predictions. One key feature of this design is that some workers are randomly informed of future piece rate and leisure opportunities a full period in advance. Previous studies on intertemporal effort allocation feature either imperfectly anticipated shocks (Camerer et al.

[8] This statement is not meant to downplay the notable concerns over external validity of laboratory experiments. For a detailed discussion, please see Charness and Kuhn (2011); Falk and Heckman (2009); Levitt and List (2007).

[9] I find a positive elasticity of effort with respect to that period's piece rate of approximately 5% to 10%. In comparison to previous papers, this is a small but significant elasticity (Chetty et al. (2011)). I also find a significant negative effect on effort when given access to cell phones in some specifications.

(1997); Oettinger (1999); Pistaferri (2003)) or fully anticipated shocks (Lozano (2011); Fehr and Goette (2007)); none (to my knowledge) intentionally manipulate the degree of anticipation of piece rate or leisure shocks.

I am able to differentiate between reciprocity and effort momentum using this variation in anticipation. First, one would expect a reciprocating worker to work harder at the time of receiving the news of an increased piece rate, not just the period in which the higher piece rate is in effect.[10] I find no such evidence. Second, for those who enjoy the additional leisure opportunity (cellphone), one would expect them to reciprocate with higher effort in surrounding periods.[11] I also uncover no evidence of this, but rather find effort is significantly reduced following phone access for those affected, as consistent with effort momentum. Thus, effort stickiness seems unlikely to be driven by reciprocity in this setting.

To address concerns about on-the-job learning, the experiments feature extensive "training" periods. Analyzing the training period data suggests that subjects reach full competency with the tasks within the first 3 minutes (see Figures 3 and 4). After this time average output is remarkably flat, as opposed to increasing output predicted by an on-the-job learning model.[12] This may not be surprising given the extreme simplicity of the tasks and is further confirmed by post-experimental surveys (see Section 3). In an experiment with multiple post-treatment periods, post-treatment effort continues to converge to baseline as predicted by momentum, rather than stay elevated as predicted by on-the-job training. Lastly, the effort stickiness implies implausibly large effects under a model of on-the-job learning, addressed in more detail in Section 6.

One might think that "switching costs" could drive this effort stickiness, but the experimental design allows me to investigate effort momentum in the absence of such switching

[10]Indeed, if one expects to find evidence of reciprocity either pre- or post-treatment, Gneezy and List (2006) suggests that the pre-treatment effects should be larger than post-treatment due to the declining effects of reciprocity over time.

[11]Some recent evidence regarding the importance of non-monetary gifts suggests this might even trigger greater reciprocation than the financial rewards (Kosfeld and Neckermann (2011); Bradler et al. (2013); Kube et al. (2012)).

[12]Even if one allows for on-the-job learning to be combined with income effects in such a way to produce flat output, one would expect to see the efficiency increases result in increased leisure time. The evidence instead suggests that leisure time is also flat or even declining for the control group.

costs. In particular, increasing the piece rate to induce greater effort would not result in switching costs as the subject would remain engaged in the task, yet there is still evidence of effort stickiness in the following period. In addition, the experimental tasks employed are able to be stopped and resumed easily as the time investment is small. Thus, to the extent that interruptions incur additional switching costs, my estimates of effort momentum could represent a lower bound of the total effort loss.[13]

Other potential explanations for effort stickiness, such as neoclassical income effects or income reference dependence, are outlined in the section on theoretical predictions and following the experimental results. In addition to not accurately matching the comparative statics found, these theories were also tested with one additional treatment involving the salience of income. I find this information salience had no effect on piece rate effects, further suggesting that effort momentum is the most parsimonious theory to explain the evidence at hand.

After addressing alternate theories, I employ an instrumental variable (IV) approach in which the previous period's piece rate and leisure options influence the previous period's effort.[14] The primary concern with this approach would be if a past period's piece rate could directly influence future effort (e.g. via a model of reciprocity). This is distinct from a model of effort momentum, where the previous period's piece rate influences this period's optimal effort only through previous period's effort.

To summarize the main findings, I find that approximately 40-50% of effort changes persist for 5 minutes even after the incentives return to baseline levels. This increase continues to decline exponentially over multiple periods. Framed another way, after an interruption of effort, it takes about 15 minutes to return to 90% of pre-interruption effort levels. Structuring the findings using this momentum parameter also provides a way to transport findings to new populations or environments (Levitt and List (2007); Falk and Heckman (2009)). For example, this estimate of 40-50% was replicated using a different "slider" task as discussed

[13]One might expect cell phones to incur greater switching costs as there is a change in user focus. This might explain why I find larger estimates of effort stickiness for cell phones in some specifications.

[14]Using asymptotics to remove the known bias, I also employ a Panel Median Unbiased Estimator as an alternate specification and find very consistent results. See Appendix Section 10.6.

in Appendix 9.3.

To address whether this productivity loss can be prevented with knowledge, I treat some subjects with information about future piece rates and leisure opportunities. Analysis shows this advance information does not impact productivity. This suggests the average subject follows a "naive" model of momentum as opposed to a more "sophisticated" model. These models are discussed in more detail in Section 2.

To put these findings in context, research suggest US knowledge workers are interrupted somewhere between 12 and 40 times a day depending on work environment. Given the ubiquity of interruptions and the large number of US knowledge workers, it is perhaps not surprising that the resulting momentum loss is quite high. If the average knowledge worker suffers 15 interruptions per work day, this will result in about 1 hour of productivity loss due to momentum alone.[15] This works out to 200 hours per year per full-time worker. If each knowledge worker earns an average of \$21 per hour, then 56 million workers[16] would lose \$235 billion per year from momentum loss alone.[17] One important caveat is that if the reduced productivity results in greater leisure, this figure would also not account for any welfare gains from this leisure – however, workers tend to report interruptions as a major source of stress in the workplace, making welfare gains unlikely (Mark et al. (2008)).[18] While there are serious concerns about generalizing evidence from students,[19] this back of

[15]Momentum loss might also explain why subjective reports of time wasted due to interruptions (often as high as 40% of total work time) tend to be higher than the observed time loss (roughly 20% of total work time).

[16]From 2014 Current Population Survey, number of management, professional, and related workers. Most common examples include software developers, financial managers, accountants, lawyers, school teachers, registered nurses, and chief executives. This number also corresponds with the 77 million workers that reported using a computer at work in 2003 by the Bureau of Labor Statistics.

[17]Though this is just a rough estimate for a number of reasons. One might expect that the individuals who earn higher than average wages are less prone to interruptions or momentum loss. Alternatively, perhaps wages have already been lowered to account for interruption loss, underestimating the true value of productivity loss.

[18]This may not be surprising given that interruptions are unplanned, making such 'breaks' in effort unlikely to be ex ante optimal from the interrupted worker's perspective. Thus, even if the productivity loss following interruption increases utility through leisure, it may be used as a substitute for a more relaxing (planned) break. Thus, the utility from such leisure could be a net welfare loss as it disrupts the optimal on-the-job leisure schedule.

[19]For example, students may lack the workplace experience that could help reduce momentum loss. On the other hand, one might also expect college students to be better than average at avoiding momentum loss as they have passed college admissions. In addition, while the environment studied is similar to what many

the envelope calculation demonstrates the potential value of further research.

Additional Contributions to Literature

In addition to the broader intertemporal labor supply literature, this paper builds on an extensive literature that uses laboratory experiments to investigate labor economic theories (for a review see Charness and Kuhn (2011)). While many papers in this literature have intertemporal implications (e.g. Rabin (1993); Dickinson (1999); Gneezy and List (2006); Levitt and List (2007); Buser and Peter (2012); Kube et al. (2012); Milkman et al. (2013); Kessler and Norton (2015)), I believe this is the first laboratory study to vary incentives over short time periods specifically to investigate intertemporal spillovers.[20]

As I find evidence of effort momentum over short time periods, this paper also suggests not to estimate individual fixed effects with short time panels. This is discussed in more depth within Section 4, but follows from earlier work on the asymptotic bias from fixed effects in time recursive models, proven in Nickell (1981). Although this has been noted when estimating effects of training programs (Card and Sullivan (1988)), this study presents new evidence that the bias may be present in more general labor settings.

This new evidence of effort momentum may also provide new interpretations of existing labor studies. While pursuing other research topics, a few[21] recent studies have uncovered intertemporal evidence consistent with momentum. In Cardella and Depew (2015), experimental subjects stuff fewer envelopes after being quantity constrained in the first period (compared to control). By itself, however, this could be evidence of on-the-job learning or reduced reciprocity due to constrained output. Bradler et al. (2015) experimentally varies payment structures in one period and also finds some persistence in effort after those incentives have been removed. For example, those who face a tournament structure exert greater effort for both creative and uncreative tasks, which significantly persists in the following

knowledge workers face, the tasks employed differ, raising additional concerns about external validity.

[20]Following Corgnet et al. (2014), it is also among the first to experimentally vary leisure opportunities within the laboratory, providing additional evidence on the effect of leisure on effort (Chapela (2007); Connolly (2008); Lozano (2011); Ward (2012)).

[21]This literature review is unlikely to be comprehensive as previous studies may have suffered from bias driven by individual fixed effects or may have simply omitted reporting intertemporal spillovers.

period. Yet this effect was strongest among tournament winners, making it theoretically unclear whether there was a "joy of winning" effect as in Kräkel (2008) or whether tournament winners, who worked hardest, simply had the largest spillover effects. Despite these confounds, this suggests that effort momentum might fill a gap between theory and empirics that has previously gone unreported.

The burgeoning literature on multitasking within economics may also benefit from study of effort momentum. Buser and Peter (2012) employ a real-effort experiment and find that subjects forced to work sequentially were more productive than subjects forced to work on tasks simultaneously. Additionally, workers allowed to work sequentially or simultaneously were also less productive than workers forced to work sequentially – mirroring the "naive" theory of effort momentum. However as Coviello et al. (2014) outlines theoretically, this sort of task-juggling may provide incentives to work harder when effort cannot be observed (as others compete for the worker's attention).

Even though my time scale is short, understanding intertemporal labor supply has important implications for labor markets and public policy. For example, if the intertemporal substitution elasticity is large and positive, one might interpret the lower pay of "flexible" positions as resulting from compensating differentials (Goldin (2014)) or a "Rat Race" equilibrium[22] (Akerlof (1976); Landers et al. (1996)). These outcomes might invite labor market policies to increase total surplus.[23] On the other hand, if this elasticity is small or negative, then the documented wage-flexibility tradeoff may be driven by firms' production and cost functions.[24] In this case, labor restrictions on hours could reduce firm efficiency.

[22] By a rat race equilibrium, I mean one in which workers work inefficient hours or effort to signal hard to observe qualities (such as ability) to employers. This was first proposed in a theoretical framework by Akerlof (1976), and there has been evidence to suggest this occurs in at least law firms (Landers et al. (1996)). As Arulampalam et al. (2007) point out, this rat race equilibrium could contribute to gender pay gaps, especially toward the top.

[23] For example, Goldin (2014) calls for "alterations in the labor market, in particular changing how jobs are structured and remunerated to enhance temporal flexibility" to reduce gender inequality in labor markets. Generally, if firms have imperfect information about worker productivity, workers may be afraid to express a desire for flexibility even though such a change would increase total surplus for the worker and firm. For example, if an hour's potential productivity is correlated with leisure opportunities, a worker's desire for flexibility could signal a desire to only work low productivity hours.

[24] For example, this could result either from per employee fixed costs (requisite search and training, benefits, or capital) or from increasing returns to hours worked (increasing worker knowledge flows, being available for clients).

The experimental design I constructed provides additional evidence on this intertemporal elasticity and is the first within this literature to vary anticipation of piece rates. As Fehr and Goette (2007) stress, the anticipation of wage changes is critical to interpretation of these elasticities.[25] Yet, informing workers about wage changes has the potential to trigger reciprocity toward the employer (Rabin (1993); Fehr and Schmidt (2006)). This complicates the interpretation of previously measured intertemporal elasticities, as anticipation and reciprocation are linked in studies with anticipated wage changes.[26] Furthermore, Gneezy and List (2006) suggest reciprocity may decline over time, potentially introducing an upward bias to wage elasticities measured over a short time period (if reciprocity is a large factor).[27]

The experimental design also allows me to address whether higher piece rates induce reciprocity. While the role of reciprocity in labor markets is an area of active research (for review see Kessler (2013); Levitt and Neckermann (2014)), I believe this is the first paper to tackle this particular question. The answer is ex ante unclear because while a higher piece rate expands the budget set, the worker must still exert effort to receive the benefits. Most previous studies testing for reciprocity in labor markets employ flat hourly wage variation in a reputation free environment (Kube et al. (2012); Fehr et al. (2008); Englmaier and Leider (2010, 2012); Kessler (2013); Gneezy and List (2006); Charness (2004)). As workers have arguably no financial incentive to work harder, evidence of greater effort is taken as evidence of reciprocity. Recent work such as Kube et al. (2012); Bradler and Neckermann (2015) suggests workers may reciprocate based on their impressions of employer intentions, rather than the actual "gift".[28] While my study is more suggestive on this point, I find no evidence that additional leisure opportunities induce reciprocity.

[25]In studies with imperfect knowledge of future wage variation, workers with better predictions may react differently to wage changes and this could also be correlated with ability. Pistaferri (2003) attempts to overcome this issue by incorporating elicited worker expectations into estimates and finds a larger EIS estimate. Unfortunately, output or effort levels are not directly observable in their dataset.

[26]This "total effect" may be the most policy relevant measurement, but as Kube et al. (2012) have found, the framing of the wage increase can contribute greatly to the degree of reciprocation. Thus, there may be no unified policy relevant reciprocation measurement if it varies greatly based on implementation.

[27]Though, as authors admit, whether the reciprocity effects would return on the second day of work is an important open question for interpretation.

[28]However earlier work from Charness (2004) suggests that exogenously determined wages elicit almost as much reciprocation as employer designated wages.

I also address whether salience of information induces workers to engage in income targeting. In a real-effort laboratory experiment, Abeler et al. (2011) find workers exert more effort when facing a chance of a higher fixed payment. Pope and Schweitzer (2011) finds evidence of loss aversion in a high stakes labor market (professional sports). In these settings, the reference point is at least partially induced by the environment (i.e. the magnitude of the outside option in Abeler et al. (2011) and golf par score in Pope and Schweitzer (2011)), but there remains some uncertainty whether information about own performance can alter endogenously chosen reference points. To investigate this possibility, I vary whether the worker sees her total earnings or her past period earnings and find it does not alter effort allocation.

The remainder of the paper is organized as follows. Section 2 derives straightforward comparative statics to distinguish the theories suggested above. Section 3 outlines the experiment designs. Section 4 discusses the specifics of the estimation strategy. Section 5 presents the results. Section 6 addresses additional concerns of alternate theories and Section 7 concludes.

1.2. Predictions

In this section, I derive predicted changes in labor supply to inform the experimental designs. I discuss three model classes below: (i) (neoclassical) time separable utility, (ii) effort momentum, and (iii) reciprocity. Additional model discussion may be found in Section 6 and the Appendix. I find straightforward comparative statics that can then be tested by the experimental design presented in Section 3.

1.2.1. Time Separable Utility

To serve as a starting point for predictions, I present a time separable model in which an agent maximizes lifetime utility

$$U_0 = \sum_{t=0}^{T} \delta^t u(c_t, e_t, \gamma_t)$$

where $\delta < 1$ represents the discount factor, $u(\cdot)$ represents the one-period utility function, c_t represents consumption, e_t is effort, and γ_t is a taste shifter that alters preferences for working in particular time periods. In my setting, γ_t can incorporate the varying leisure opportunities available, such as cell phone access. I further assume that the utility function is differentiable and $u_c > 0$, $u_e < 0$ and strictly concave in c_t and e_t. The lifetime budget constraint is given by

$$\sum_{t=0}^{T} p_t c_t (1+r)^{-t} \quad \leq \quad \sum_{t=0}^{T} (w_t e_t + y_t)(1+r)^{-t}$$

where p_t represents prices at time t, w_t the piece rate at time t for each unit of effort e_t, and y_t represents non-labor income. Also the interest rate r is assumed to be constant, but this does not impact the sign of the comparative statics.

As shown in Fehr and Goette (2007), along the optimal path, this model can be equivalently represented as an individual optimizing a static one period utility function that is linear in income. This can be written as:[29]

$$v(e_t, \gamma_t) \quad = \quad \lambda w_t e_t - g(e_t, \gamma_t)$$

where $g(e_t, \gamma_t)$ is strictly convex in e_t and captures the discounted disutility of effort. λ captures the marginal utility of life-time wealth. In this formulation, $\lambda w_t e_t$ represents the discounted utility from total income earned in period t.

Thus, as w_t increases, the optimal e_t^* will also increase. The effort exerted today is only influenced by past piece rates through the marginal utility of life-time wealth λ. In the literature on measuring temporary wage or piece rate shocks, this λ is assumed constant as the total impact on lifetime wealth is very small, implying small changes in λ (Fehr and Goette (2007)). Therefore, with no income effects, a single period's piece rate would have no impact on effort in surrounding periods.

[29]Note that this formulation omits the price path $\{p_t\}$ and δ as these are not the objects of study. If these elements change, the corresponding g function would also change, but it could still be written in a similar format.

If one allows for income effects, additional income would increase the attractiveness of leisure given the concavity of consumption. As a result, allowing for income effects would reduce effort in periods surrounding a piece rate increase.[30]

Lastly, if one allows for leisure technology γ_t to increase the disutility of effort (e.g. harder to work when the World Cup is on), then increasing leisure technology would decrease optimal effort e_t^* in that period. As with piece rates, in the absence of income effects there are no predicted spillovers on the surrounding periods. If one allows for income effects, then an agent would work harder in surrounding periods (say before or after the World Cup game). This follows as the reduced lifetime income (from the high leisure time) would increase the marginal utility of lifetime income, λ.

1.2.2. Effort Momentum

Effort momentum is a model in which past period's effort directly influences the disutility of future periods. For example, working hard may engage a flow-like state in which future effort is less costly.[31] Alternatively, if effort is interrupted for a period ($e_t = 0$), the worker may face greater disutility to start working again.[32]

To capture these ideas, I present a model in which an agent encounters lifetime utility:

$$U_M = \sum_{t=1}^{T} \delta^{t-1} u(c_t,\, e_t,\, e_{t-1},\, \gamma_t)$$

where $\delta < 1$ represents the discount factor, $u(\cdot)$ represents the one-period utility function, c_t represents consumption, e_t is contemporaneous effort, and γ_t is a taste shifter that alters preferences for effort in particular time periods. In my setting, γ_t can incorporate varying leisure opportunities available and will be referred to as leisure technology.

[30] Under the current experimental design, this would have very similar predictions to a model in which the agent has a daily income target.

[31] This is consistent with an extensive psychology literature on "flow", cf. Nakamura and Csikszentmihalyi (2002); Schaffer (2013), in which workers enter a state where the disutility of work is reduced as subjects report losing a sense of self.

[32] It is worth noting that these intertemporal effects do not necessarily have to be positive – a conceptually similar model proposed by Fehr and Goette (2007) includes a cost function in which greater effort today increases the marginal disutility of effort in the next period, perhaps due to stress or physical exertion.

I further assume that the utility function is twice-differentiable in its arguments with $u_1 \geq 0$, $u_2 \leq 0$, has a positive cross partial $u_{23} \geq 0$. With these assumptions, consumption is enjoyable, effort is unenjoyable, and past effort decreases the marginal disutility of effort. I also assume that leisure technology makes effort more costly in utility terms ($u_{24} \leq 0$), but also has no positive effect on consumption ($u_{14} \leq 0$).[33] Lastly, that effort does not make consumption more enjoyable ($u_{12} \leq 0$, $u_{13} \leq 0$). The lifetime budget constraint is given by

$$\sum_{t=1}^{T} p_t c_t (1+r)^{-t} \quad \leq \quad \sum_{t=1}^{T} (w_t e_t + y_t)(1+r)^{-t}$$

where p_t represents prices at time t, w_t the piece rate at time t for each unit of effort e_t, y_t represents non-labor income, and r is the interest rate from one period to the next.[34] I also assume there is no change in lifetime marginal utility of wealth λ is constant, as the total impact on lifetime wealth is very small. This is in line with other field and laboratory experiments in the labor economics literature (Fehr and Goette (2007); Camerer et al. (1997)).

Sophisticated Momentum

Sophisticated Momentum is the model as described above, in which the agent correctly realizes that today's effort will influence tomorrow's marginal disutility of effort.

Proposition 1.2.1 *Under the above assumptions, effort is monotonic non-decreasing in past, present, and future piece rates. Alternatively, effort is monotonic non-increasing in past, present, and future leisure technology expansions.*

Proof Application of supermodularity theorems. See Appendix Section 10.4. ■

The intuition for these comparative statics is straightforward. If a worker is aware that effort now will decrease the cost of effort in the next period, then the periods' optimal efforts will

[33] Otherwise, increased leisure technology could boost desire for consumption to the extent that the agent works more to increase lifetime wealth.

[34] While the interest rate r is assumed to be constant for notation simplicity, this assumption does not impact the sign of the comparative statics.

move together due to the spillover. For example, if the worker faces a higher piece rate next period, then next period's effort will become marginally more valuable. As work in the present reduces the costs of working next period, the marginal benefit of working in the present has also increased. By a similar argument, if the worker faces greater leisure opportunities next period, the benefits (due to effort momentum) of working today has also decreased.

Naive Momentum

Although the agent experiences the effects of momentum, it may be possible that either the agent does not realize this momentum will occur in the future, or otherwise uses an exogenous reference for future effort.[35] I call this model Naive Momentum. In this model, at period t, the agent maximizes a discounted stream of future utility:

$$U_t = u(c_t, e_t, e_{t-1}, \gamma_t) + \sum_{j=t+1}^{T} \delta^j v(c_j, e_j, \gamma_j)$$

and will formulate plans of this period and future period's effort. Note that the $v(\cdot)$ function above does not have e_{t-1} in its arguments. However, once the agent actually arrives at time $t+1$, he correctly incorporates previous period's effort into his lifetime utility:

$$U_{t+1} = u(c_{t+1}, e_{t+1}, e_t, \gamma_{t+1}) + \sum_{j=t+2}^{T} \delta^j v(c_j, e_j, \gamma_j)$$

This will cause the agent to revise his plans he made in time period t. The agent also faces the same budget constraint as before:

$$\sum_{t=0}^{T} p_t c_t (1+r)^{-t} \leq \sum_{t=0}^{T} (w_t e_t + y_t)(1+r)^{-t}$$

Proposition 1.2.2 *Under the assumptions above, there is an equivalent period utility func-*

[35] One possible justification for this is the literature on Projection Bias, see Loewenstein et al. (2003); Conlin et al. (2007); Simonsohn (2010). Under such projection bias, a tired individual may incorrectly project that they will always be tired – but if he started working harder he may be surprised to find he isn't as tired as expected.

tion

$$u = \lambda w_t e_t - g(e_t, e_{t-1}, \gamma_t)$$

This form demonstrates that effort is increasing in past and present piece rates, but future piece rates have no impact. By the same token, effort is decreasing in past and present leisure technology, but future leisure technology has no impact.

Proof Proof of the $g(\cdot)$ function equivalence and its convexity is provided in Appendix 9.1, but builds on work by Browning et al. (1985) and Fehr and Goette (2007). A brief proof for the comparative statics is provided below.

Consider the effect of an increase in w_{t+j}. In the first period, the first order condition states:

$$g_e(e_1^*,\, e_0,\, \gamma_1) \quad = \quad \lambda w_1$$

e_0 cannot be influenced by any $w_{t'}$ by construction, as time period 0 is before any information is received. γ_1 are not choice variables, they are only exogenously given. Thus when I take the derivative with respect to w_{t+j} to get:

$$g_{ee}\frac{de_1^*}{dw_{t+j}} \quad = \quad 0$$

Which, as $g_{ee} > 0$ gives us the effect in the first period of 0. In time period t, to complete the induction proof I assume $\frac{de_{t-1}^*}{dw_{t+j}} = 0$ and look to prove the same is true for $\frac{de_t^*}{dw_{t+j}}$. This follows from taking the total differential of the first order condition:

$$\begin{aligned} g_{ee}\frac{de_t^*}{dw_{t+j}} + g_{e2}\frac{de_{t-1}}{dw_{t+j}} &= 0 \\ \Rightarrow \frac{de_t^*}{dw_{t+j}} &= 0. \end{aligned}$$

Thus, by induction, optimal effort prior to a piece rate increase is unchanged when holding λ constant. This follows from the assumption of naivety that the agent does not anticipate future momentum. However, once the agent reaches the period with higher piece rates, an increase in the piece rate still elicits greater effort:

$$\frac{de_t^*}{dw_t} = \frac{\lambda}{g_{ee}} > 0$$

This follows from the convexity of g w.r.t. e_t^*. The same sign can be seen by looking at the total derivative with respect to past piece rates, w_{t-1}:

$$\frac{de_t^*}{dw_{t-1}} = -\frac{g_{e2}}{g_{ee}} \frac{de_{t-1}}{dw_{t-1}} > 0$$

As $g_{ee} > 0$, $\frac{de_{t-1}}{dw_{t-1}} > 0$ and $g_{e2} < 0$ (as $u_{23} > 0$). The proofs for leisure technology are the same as above with opposite signs (as leisure technology makes effort more costly, rather than less). ■

Although the above proposition gives us the required comparative statics of interest for naive momentum, considerably more can be said with an additional restriction on the period utility function. Without assuming a specific functional form, one can show that that the optimal effort will follow a linear time recursive structure.

Proposition 1.2.3 *Assuming further that* $u(c_t, e_t, e_{t-1}, \gamma_t) = q(c_t, e_t - \rho \cdot e_{t-1}, \gamma_t)$ *with* $|\rho| < 1$, *then optimal effort will follow a time recursive structure*

$$e_t^* = \rho \cdot e_{t-1} + z(w_t, \gamma_t)$$

with $z(\cdot)$ *increasing in* w_t *and decreasing in* γ_t.

Proof By a similar proof as above, the FOC will be

$$\begin{aligned} -q_e(c_t^*, e_t^* - \rho e_{t-1}, \gamma_t) &= \lambda w_t \\ q_c(c_t^*, e_t^* - \rho e_{t-1}, \gamma_t) &= \lambda p_t \end{aligned}$$

As q is strictly concave over the first argument, this allows for inverse of q_c:

$$c_t^* = q_c^{-1}(\lambda p_t, e_t^* - \rho e_{t-1}, \gamma_t)$$

Which can be inserted into the first FOC to give:

$$-q_e(q_c^{-1}(\lambda p_t, e_t^* - \rho e_{t-1}, \gamma_t), e_t^* - \rho e_{t-1}, \gamma_t) = \lambda w_t$$

Thus, a new utility function $\lambda w_t e_t - h(e_t^* - \rho e_{t-1}, \gamma_t)$. The convexity of $h(\cdot)$ gives us an inverse function for h_1:

$$e_t^* = \rho e_{t-1} + h_1^{-1}(\lambda w_t, \gamma_t)$$

As this is a special case of the first proposition (if $\rho > 0$), optimal effort e_t^* will still be an increasing function of w_t and decreasing in γ_t. In addition, past effort positively influences current effort and future piece rates or leisure technology does not influence current effort.■

1.2.3. Reciprocity

Consider instead a model in which changes in w_t and γ_t induce a desire to reciprocate. As formulated, this is similar to the time separable utility, but with an additional component of utility based on the piece rates and leisure offered across all time periods:

$$U_R = \sum_{t=0}^{T} \delta^t u(c_t, e_t, \gamma_t) + \alpha(\{w_t\}, \{\gamma_t\}) \cdot \left(\sum_{t=0}^{T} \delta^t e_t \right)$$

In which u has the same properties as outlined above ($u_1 \geq 0$, $u_{11} < 0$, $u_2 \leq 0$, $u_{22} < 0$, $u_{23} < 0$) and with $\alpha(\cdot)$ strictly increasing in its arguments. In this model, increases in future or past piece rates can increase the marginal utility of effort through the 'altruism' or 'fairness' function α. This model is similar to ones found in Rabin (1993); Fehr and Schmidt (2006). Extending the work of Browning et al. (1985), this utility function can be reformulated as a series of period utility functions:

$$v(e_t) = [\lambda w_t + \alpha(\{w_t\}, \{\gamma_t\})] \cdot e_t - g(e_t, \gamma_t)$$

In a simple two period model for illustrative purposes, the agent receives additional marginal utility based on w_1 and w_2. For simplicity, I assume that this additional utility is linear in piece rate and effort, $\alpha(\{w_t\}, \{\gamma_t\}) = \alpha_1(w_1 + w_2) + \alpha_2(\gamma_1 + \gamma_2)$. Thus the agent is maximizing:

$$\begin{aligned}
U_R &= v(e_1) + v(e_2) \\
v(e_1) &\equiv (\lambda w_1 + \alpha_1 w_1 + \alpha_1 w_2 + \alpha_2\gamma_1 + \alpha_2\gamma_2)e_1 - g(e_1, \gamma_1) \\
v(e_2) &\equiv (\lambda w_2 + \alpha_1 w_1 + \alpha_1 w_2 + \alpha_2\gamma_1 + \alpha_2\gamma_2)e_2 - g(e_2, \gamma_2)
\end{aligned}$$

In this setting, increasing the piece rate can increase reciprocity, even in surrounding periods. Under this simple model, if $\alpha_2 > 0$ and e_1, e_2 is an interior solution, then $\frac{\partial e_2}{\partial \gamma_1} > 0$. Likewise, if $\alpha_1 > 0$ and e_1, e_2 is an interior solution, then $\frac{\partial e_2}{\partial w_1} > 0$. Similar intuitions apply for future piece rates or leisure technologies when informed in advance. For proofs, please see Appendix Section 10.4.

1.2.4. Summary

Owing to space limitations, several theories have been moved to a discussion following the results. To summarize the most relevant theories, I present the following table that outlines how effort at time t will respond to piece rates and leisure technologies at different times (past, present, and future):

Please note that not all reference models give precise comparative statics in some situations, as shown in Brandon et al. (2014). The comparative statics for reference models above are under the case that the reference or target is strong and influences the intertemporal effort allocation. For example, an income target of \$1000 would not be possible to achieve in a laboratory setting, but would also not influence the intertemporal results (the agent would appear as a neoclassical time-separable agent). More adaptive models of reference dependence are discussed in Section 6.

Table 1: Predictions Summary Table

Models		Effort at time t in response to increase in: Piece Rate at time: $t-1$	t	$t+1$	Leisure Tech at time: $t-1$	t	$t+1$
Time Separable	No Income Effects	0	+	0	0	-	0
	Income Effects	-	+/-	-	+	-/+	+
Momentum	Naive	+	+	0	-	-	0
	Sophisticated	+	+	+	-	-	-
Reciprocity		+	+	+	+	-	+
On-the-job Learning		+	+	+	-	-	-
Income References	Period Target	0	-	0	0	-	0
	Total Target	-	+/-	-	+	-	+
	Previous Period	+	-	0	-	-	0
Experiment Results*		+	+	0	- or 0	- or 0	0

(*see Section 5 for details)

Also please note that although On-the-job Learning and Sophisticated Momentum have the same predictions for the 6 comparative statics above, there are additional tests to distinguish these two hypotheses. For example, if the gains are primarily driven by learning, one would expect either (a) increasing quantity over time or (b) increasing leisure engagement over time. Neither of these are found to occur. There are also reasons to believe that the magnitudes involved make learning a very unlikely possibility, see Section 6 for more details. In practice, as I find evidence for Naive Momentum, these comparisons are not as crucial.

1.3. Experiment Design Overview

In order to test these comparative statics, I investigate how effort responds to changes in (i) past (ii) contemporaneous and (iii) future piece rates and leisure opportunities. These correspond to the 6 columns of the Prediction Summary Table above. These hypotheses were tested over two similar experiments (differences outlined below).

In both experiments, subjects complete incentivized real-effort tasks in a laboratory setting. The tasks involve counting images and are performed on a computer. This is similar to previous labor economics experiments studying effort in the laboratory, especially Abeler et al. (2011). Subjects count particular images from a matrix of 98 images, as can be seen in Figure 1.[36] This task was selected as it requires little to no training, but is menial and requires effort.[37] In post experiment surveys, subjects often mention the task is boring (see Appendix Figure 1), in line with findings presented in Abeler et al. (2011). Thus the primary measure of effort is the number of problems solved correctly – consistent with the experimental labor literature (Charness and Kuhn (2011); Fehr and Goette (2007)).[38]

In line with Corgnet et al. (2014); Eriksson et al. (2009); Charness et al. (2010) and to mirror many labor contexts outside of the laboratory, I introduce a baseline leisure activity. Specifically, the participants were allowed to watch YouTube.com videos at any time instead of performing counting tasks (see bottom of Figure 1). To help make YouTube videos a potentially worthwhile leisure activity, a pair of headphones was attached to every computer. However, as the video was located below the counting problem, it was difficult to engage in both simultaneously. In the appendix, I confirm that YouTube videos were indeed a time substitute for effort.

As discussed in Section 2, many models of effort allocation allow for changes in either piece rates or leisure options to impact effort. To test these models, I experimentally varied the piece rate and leisure opportunities in specific periods. Although the piece rate varied in some periods, every period contributed to final earnings. This was done to focus on intertemporal substitution as opposed to regret or risk aversion. Paying in every period also allowed me to distinguish between potential "daily" income targeting and "period" income targeting models. Final payment also included a flat $10 participation fee so long

[36]Abeler et al. (2011) has agents counting zeroes in a string of 100 numbers. This exact task was not feasible in a web browser with a "search" feature, which makes the task trivial as one can merely search for 0. As a result, I ask the worker to count either heart or drop icons (randomized at the subject level). Only one subject tried bypass the task by searching the "source code" (after being asked not to) and is dropped from analysis.

[37]Gill and Prowse (2012) employ a task with sliders that also has attractive properties (further outlined in Gill and Prowse (2011)) – this task was employed in a replication experiment with very similar results, see Appendix 9.3. However, focusing on the task similar to Abeler et al. (2011) also allows for a closer comparison to their results, including testing for possibility of reference dependence.

[38]Though output and effort may not be perfectly correlated, changes in the production function are unlikely to explain evidence provided, as discussed in a Section 6.

as they followed laboratory guidelines (e.g. no food, no talking).[39] However no payments were made until the end of the entire session.[40]

To vary the leisure opportunities, some subjects were randomly assigned access to their cell phones. The laboratory employed for this study, Wharton Behavioral Lab, ordinarily has a strict no phone policy to improve study compliance and concentration. This policy was put in place because participants have a tendency to want to text, browse the web, and play games on their cell phones during the lab session. Thus, phone access has the potential to represent an increase in the marginal utility of leisure (γ_t from Section 2).[41] This is conceptually similar to experiments conducted in Corgnet et al. (2014) which allowed some users to browse the internet to expand possible leisure activities participants face.[42]

Prior to being allowed to start each period, the subjects had to correctly answer questions about the upcoming period's piece rate and cell access. These procedures were implemented to ensure subjects fully understood the incentives they faced.[43] In addition, counters at the bottom kept track of current earnings (as in Abeler et al. (2011)) as well as visual indications for whether phone use was permitted.[44] In post experiment surveys, 98% of

[39]It was made clear and reiterated that they did not need to solve any problems to guarantee their $10 participation fee. In practice, every participant adequately followed the laboratory guidelines and received the $10 participation fee.

[40]Paying at the end was both a practical necessity given length of the periods and also mirrors the design of Fehr and Goette (2007).

[41]The subjects of the experiment were University of Pennsylvania undergraduates. The second experiment surveyed cellphone access – only 8 out of 422 subjects (2%) did not bring a cellphone to the laboratory. Even though phone quality may vary or some subjects may not have a phone, this will not impact estimate validity if randomization was adequately done. However, this research will only be able to answer whether access to phones already owned by the subjects influence effort rather than the effect of access to a particular phone. This was done in part because introducing a new cell phone would lead to significant learning, additional experimental cost, and may not represent the same expansion of leisure opportunities as if the individual owned the phone (e.g. no contacts, no texts, etc.)

[42]When given access, the students could also use the internet on their phones, so internet access could be seen in some way as a lower bound of the potential leisure opportunity faced by allowing phone use. Other leisure technology expansions were considered, but deemed too difficult to adequately monitor under the current lab setup. With cell phones, lab assistants were able to quickly verify whether cell phone users were allowed to use the cell phone at that time.

[43]In one pilot study, rather than quiz the subject on the piece rate, the website merely didn't allow them to continue until 30 seconds have passed. This allows me to investigate potential salience effects from repeating the piece rate but in the post experiment analysis did not seem to make a difference.

[44]In addition to the subject's earnings for the current period, either (i) the total earnings or (ii) previous period earnings are displayed on the screen at all times. This treatment serves as a supplementary test for income targeting and is explained in more detail in Section 6.

Table 2: Design Summary Table

	Experiment 1	Experiment 2
Task	Image Counting	Image Counting
Location	Wharton Behavioral Lab	Wharton Behavioral Lab
Subjects	155 UPenn Undergraduates	422 UPenn Undergraduates
Number of Treatment Periods*	6	3
Duration of Treatment Periods	5 minutes	5 minutes
Duration of Pre-Treatment	5 minutes	15 minutes
Baseline Piece Rate (US $)	0.05	0.05
Piece Rate Treatments (US $)	0.15 or 0.30	0.03 or 0.08 or 0.15
Baseline Leisure Access	Youtube.com	Youtube.com
Leisure Access Treatment	Phone Access	Phone Access
Advance Information Occurs	Periods 1, 3, 5	Randomly in Period 1
Instructions followed by	30 second timer and Quiz	30 second timer and Quiz
Counters at Bottom	Period Earnings and Total Earnings	Period Earnings and either Total Earnings or Last Period Earnings
On Screen Timer	No	Yes
Image Counted	Hearts	Hearts or Drops (randomized)
Randomization	Subject Level by Computer	Subject Level by Computer

subjects report that the payments and leisure opportunities available were clear.

Although the experiments followed the general design above, I outline differences in the table below and elaborate in the following sections:

*Note: Number of treatment periods is the number of all periods after the pretreatment, i.e. periods in which individuals could differ in some way. In experiment 1, every subject experienced precisely 2 of the 6 rounds had a piece rate or leisure technology that was not baseline. In experiment 2, at most 1 period had a piece rate or leisure technology that was not baseline.

1.3.1. Experiment 1 Design

At the beginning of the session, the subject was given a series of instructions and an example problem. This was followed by one 5 minute "Pre-treatment" period to become acquainted with the task. This Pre-treatment period had the same incentives for all subjects and serves as a proxy of worker ability, as will be discussed in Section 4. For each solved problem in that period, the subject is informed she will earn \$0.05. In order to discourage random guessing, there was also a penalty for wrong answers – entering the incorrect answer three times for a single problem resulted in a deduction of \$0.20, akin to Abeler et al. (2011).[45]

The participants then completed six additional periods, each 5 minutes long, with three different possible treatments:

- **Control** – Subjects receive \$0.05 per completed problem for that period.
- **High Piece Rate** – Subjects receive a higher piece rate for that period, either \$0.15 or \$0.30.
- **High Leisure Technology** – Subjects receive the ability to access their cellphones for one period but still received \$0.05 per completed problem.

The control and piece rate treatments were calibrated using a small pilot study to allow for movement in either direction, as suggested by Charness and Kuhn (2011).

Regarding the randomization, these six treatment periods are broken up into three pairs. Each pair consisted of either two periods of Control treatment; a Control treatment and a High Piece Rate treatment; or a Control treatment and High Leisure Technology treatment. Within each pair, the order of the treatments was random in order to differentiate period effects and anticipation effects. Each subject eventually receives all three treatment pairs, potentially allowing for both between and within subject analysis. This randomization was executed at the individual level by a pseudo-random number generator seeded by computer time (down to the millisecond).

To test for adequate randomization, I investigate whether pre-treatment indicators (such as gender, self-reported SAT scores, and pretreatment performance) predict the period at which the subjects faced the High Piece Rate or High Leisure treatments. As reported in Table 2A, none of these factors individually or together are predictive of the period that they receive the treatments.[46] As a result, I conclude that the treatment randomization was adequately done given the observable characteristics.

[45] On average the participants entered about 0.67 problems per period incorrectly, about 10% of total problems correctly solved per period.

[46] For regressing "High Piece Rate" treatment period # on pre-treatment variables, the F stat corresponds to a p-value of 0.29. For regressing the "High Leisure" treatment period # on pre-treatment variables, the F stat corresponds to a p-value of 0.71. Thus, for both treatments I fail to reject the hypothesis that all coefficients are zero and that none of the observable pre-treatment variables is significantly correlated with the period in which treatments occurred.

1.3.2. Experiment 2 Design

The second experiment simplifies the first by assigning each worker only a single primary treatment, over four periods rather than seven. The first "Pre-treatment" period lasted 15 minutes, was the same for all subjects and is used to generate proxies for worker ability (see Section 4). The following "treatment" periods were all 5 minutes. The first of these featured a baseline piece rate, but could (randomly) inform the subject about the next period piece rate and leisure opportunity. In the following period, the subject receives either the baseline, a piece rate treatment, or a high leisure treatment. In the final period, the subject is returned to baseline piece rate and no access to the cell phone. As in experiment 1, randomization was executed at the individual level by a pseudo-random number generator seeded by computer time (down to the millisecond).

This experiment also expands on the first one in a number of ways. First, a second piece rate treatment arm was included, in which the piece rate is decreased from \$0.05 to \$0.03 and another treatment arm randomizing "total" vs "previous period" earnings shown. Second, by randomizing the information available for all subjects, the design eliminates concerns about "odd-period" x treatment interaction effects present in the first experiment.[47] Third, by keeping each individual to a single treatment, there may be less concern that interactions between multiple treatments confound effects. This also allowed for a longer training period to further reduce concerns about on-the-job learning. Fourth, a timer was added in accordance with Abeler et al. (2011) to minimize concerns about time uncertainty driving results. Lastly, additional variables, including specific timing and phone usage, were collected and a timer was added to the post experiment survey to improve information quality.

1.4. Empirical Specifications

In the following section, the results of the experiments will be addressed, but prior to that, three important empirical notes need to be made.

1. First, the potential presence of momentum – where the previous period's effort could directly influence this period's effort – makes this a poor setting for individual fixed effects. Estimating these

[47] In the first design, being "surprised" can only happen on treatment periods 1, 3, and 5 and "advance knowledge" can only occur for periods 2, 4, and 6. Although period fixed effects are included in most specifications, if odd-periods were interacting with treatments in some other way besides knowledge (e.g. piece rate increases are more effective in the final period), then estimates from experiment 1 could be a combination of those odd-period interaction effects and the effect of advance knowledge.

individual fixed effects will lead to bias in the estimate of the momentum.[48] Nickell (1981) proves this, but the intuition is that shocks will be partially absorbed into the fixed effect estimate rather than the coefficient estimate for the previous period's effort. This is worse when there are fewer periods as there are fewer shocks to properly distinguish the coefficient estimates.

For example, assume effort follows an AR(1) process (similar to Proposition 2.3) and there is a time-constant individual fixed component

$$e_{i,t} \quad = \quad \rho \cdot e_{i,t-1} + f_i + \beta x_{i,t} + \nu_{i,t}$$

where $e_{i,t}$ is the number of problems solved by individual i at time t, ρ captures the degree of "momentum" from the previous period, f_i is the individual ability or motivation component, $x_{i,t}$ include other shifters such as piece rate or leisure technology and $\nu_{i,t}$ is an error term. Under this model, estimating individual fixed effects will introduce an asymptotic downward bias to ρ, approximately equal to $-\frac{1+\rho}{T-1}$. In my setting with $T = 3$, even if ρ was 0.5, asymptotic estimates would become indistinguishable from 0 as $N \to \infty$. This remains an issue even though the piece rate and leisure technology are randomized.[49] To be clear, this is not an issue of error terms correlated within an individual which could bias the standard errors[50] but rather a bias in the coefficient estimates themselves.

However, this was a known issue when designing the experiment and a primary justification for the pre-treatment period. This pre-treatment period can then serve as a proxy for individual ability or motivation, taking the place of f_i. To minimize risk of overfitting the data, a non-parametric approach is employed – individuals are split into five quintiles based the number of problems solved in the pre-treatment period, then each quintile receives it's own binary indicator variable.[51] The pre-treatment period is therefore omitted from the dependent variable for all specifications. In line with Card and Sullivan (1988), I also present a random effects model with identical findings in the Appendix.

2. The second note is that, although many labor studies take logarithms of dependent and in-

[48] Simply omitting the momentum term will not solve this issue in general but rather can bias other coefficients.

[49] However it may be worth noting that period fixed effects or session fixed effects will not be biased by this momentum, as the error terms are not correlated across individuals.

[50] Throughout the paper all standard errors are clustered at the subject level to reduce the influence of error terms correlated within an individual.

[51] Though in practice the results are virtually identical when using a linear and quadratic term for number of problems solved in pre-treatment.

dependent variables, my specifications are reported at the unit level of analysis. This is done for several reasons, first being that the theory of momentum in section 2.2 suggested a unit level of analysis of effort. Given the linearity of the task, one might think effort would be closely correlated with quantity, not log(quantity). Second, while not common, some individuals did opt to solve no problems in a given period, a common issue with log forms. Third, the unit level was my ex ante specification while designing the experiment and analysis, and interpretation of new p values would be problematic after analysis has already been completed.

On the other hand, a linear formulation with OLS may be considered problematic as effort shocks cannot be too negative if effort is bound at the lower end at 0. It's also hard to represent upward effort "caps" with a linear specification. That being said I employ a log(problems correct + 1) on log(piece rate) specification and find qualitatively very similar results. While these measurement issues are important, the hypotheses are tested by the qualitative signs.

3. Third, as explained in Section 3, experiment 1 had each subject being treated to all 3 possible treatments. This helped improve power of treatment effects given the smaller sample – but it runs the risk of multiple treatments interacting to confound estimates. For example, access to a cellphone following a high piece rate period could negate additional effort resulting from momentum or reciprocity. As a result, I also perform analysis on just the first treatment received (corresponding to the first 3 periods of treatment)[52] and find that it does influence the intertemporal results in some specifications. While the interaction of treatments may be interesting, this was not the primary goal of this research study. Given this and the above difficulties of within-individual analysis in this setting, the second experiment design was simplified so that each person received only one treatment. This also allowed for a longer "training" period to further ensure the results are not being driven by on-the-job learning.

1.5. Experiment Results

Within this section, experimental results are presented together, as they are overall very similar across the two experiments. I use these results, particularly the qualitative signs, to test predictions of different theories from Section 2. I begin with the contemporaneous (same period) effects and move on to intertemporal effects.

[52]Restricting analysis to only the first treatment pair is equivalent to focusing on just periods 1 and 2; however as half of those subjects received treatment in period 2, the following period (for intertemporal analysis) would be period 3. Results change very little when limiting it to individuals who have only received baseline piece rate (0.05) and leisure (YouTube) in period 3.

1.5.1. Contemporaneous Effects

The first question is whether the primary treatments impacted contemporaneous effort as predicted by most theories of intertemporal labor supply. Recall that the primary treatments (piece rate or leisure technology) were in place for only 1 period, so this is asking whether or not effort was influenced in that treated period. This is important because if there is no effect on effort in the treated period, it would be difficult to understand why they would affect earlier or later periods.[53]

Result 1.5.1 *In accordance with most intertemporal theories of effort allocation, an increase in the piece rate significantly raised effort in the effected period. Likewise, in some specifications, there was a significant decrease in effort when subjects are offered access to their cellphones. See Tables 3A and 3B for details.*

In the first experiment, the average worker solves 0.20 to 0.45 more problems ($p < 0.01$) when faced with a 10 cent increase in the piece rate, as seen in Table 3A. This treatment estimate corresponds roughly to an effort elasticity of 2% = (0.325/7.85) / (0.10/0.05). Thus, increasing the piece rate by 50% would increase average effort in this context by approximately 1%. This elasticity is small relative to previous findings in the literature, though still significant (Card (1991); Chetty et al. (2011); Fehr and Goette (2007)). Compared to the existing literature, this low result may be best explained by an effort ceiling.[54] In other words, at a $0.15 piece rate, agents may have already been exerting close to their maximum potential effort. There is some evidence for this, as the $0.15 and $0.30 piece rates both elicited greater effort, did not significantly differ from one another. This in turn would push down the average elasticity. To examine this possibility, experiment 2 features a $0.08 piece rate (1.6x baseline) and a $0.03 piece rate (0.6x baseline) instead of the $0.30 piece rate (6x baseline). Alternatively, the low elasticity could be the result of multiple treatment interactions. As seen in Table 4A, limiting the analysis to the first treatment increases the elasticity up to about 5%. Given these issues, I believe experiment 2 represents a better estimate for the contemporaneous elasticity.

[53] That is not to say that certain combinations of theories could not predict such a finding, e.g. if an agent had a period income target but also experienced reciprocity, the two effects might cancel out in the effected period but could influence outside periods. However, given the extensive literature on piece rates influencing effort in the given period, such a null finding would likely indicate the treatment or sample size is too small (Levitt and Neckermann (2014)).

[54] Though it may be difficult to compare as previous literature often focuses on hour or participation elasticity rather than effort.

In the second experiment, I find a larger effect of a higher piece rate on effort. As can be seen in Table 3B, every 10 cents (200%) increase (decrease) in piece rate increases (decreases) the number of correct problems by 1.24 - 1.57 (17-22%). Thus, the estimated elasticity of effort with respect to contemporaneous piece rate is 9.3% for experiment 2, quite a bit higher than the 5% found in experiment 1. Although not my primary inquiry, there did not seem to be a difference in magnitude for piece rate increases compared to decreases (no significant "kink" in the slope).

As discussed in Section 2, many intertemporal labor models also predict that increasing the marginal utility of leisure detracts from effort provision. One way to test this hypothesis is by increasing the leisure options available to the subject. To the extent that these leisurely options are complements with leisure time, one would expect an increase in leisure time and a corresponding decrease in total effort. In this experiment, the agents had access to Youtube.com videos throughout the experiment, but during the "High Leisure Technology" treatment, were also given access to their cellphones. When faced with this phone access, experiment 1 subjects complete 0.43 fewer problems on average ($p < 0.05$), as seen in Table 3A. This provides support for the hypothesis that leisure opportunities can reduce effort allocation.[55] However, once the sample is restricted to the first 3 treatment periods (to eliminate potential multiple-treatment interactions), this coefficient is no longer significant (see Table 4A specifications 4 through 6). Thus it is possible then that the original effort decrease due to cellphones was driven by subjects who received access to cellphones after the increased piece rate.[56]

In the second experiment, there was no significant decrease when cell phones were permitted. This matches the finding in the first experiment once restricted to the first treatment. However, when broken down by gender, cellphones appear to reduce effort in the contemporaneous period for males, as can be seen in Tables 5A and 5B.

These contemporaneous estimates also serve to test the "period income" reference dependence model. In this model, an agent receives relatively greater disutility if she falls short of a particular income in a given period. For example, a subject might try to earn $0.50 each period and then spent the rest of the time watching YouTube videos. Under this model one would usually[57] expect to see a reduction

[55] Unfortunately the binary nature of phone access does not allow for an accurate elasticity measurement of effort with respect to leisure opportunities. However, calculations of implicit time value of the leisure opportunity might allow one derive an estimate.

[56] Alternatively this evidence might suggest that cell phones were more effective in reducing effort during the later periods, or may be due to lower power from a smaller sample sample size. Results from experiment 2 suggests the latter hypothesis, as subjects in experiment 2 are treated with cellphone access quite "late" in the session, yet the treatment does not seem to influence effort.

[57] As discussed in Brandon et al. (2014), if the piece rate increase is large enough or if the target is too large, contemporaneous effort could still increase with piece rates akin to a neoclassical time separable utility

in effort when faced with a higher piece rate (as it has become easier to earn the target income for that period). Yet, as discussed, the findings suggest the opposite direction, with an increased piece rate inducing greater effort in the period it was enacted. Therefore, the contemporaneous evidence does not support a "period income" reference point.

1.5.2. Intertemporal Treatment Effects

In addition to a contemporaneous treatment effect, pre-and post-treatment effects are important to differentiate the theories outlined in Section 2. For example, if workers followed a neoclassical time separable utility function, then as total impact on income is small, one would not expect to see any reduction or increase in effort in the periods surrounding the high piece rate or high leisure treatments.[58] Instead, I find significant stickiness in effort:

Result 1.5.2 *In the period following an increase in the piece rate, effort was also significantly higher. This is consistent with models of Effort Momentum as well as Reciprocity. See Table 4A and 4B for details. However (randomized) advance knowledge of higher piecerates did not significantly influence effort. Of the models outlined in Section 2, these results are only consistent with a model of Naive Momentum. See Table 6A and 6B for details.*

In the second experiment, the intertemporal treatment effects are quite striking, presented in Table 4B. An increase in the previous period's piece rate of \$0.10 (200%) significantly increases the effort in the following period by about 0.75 problems (10%). Thus, for experiment 2 the intertemporal elasticity is about half of the contemporaneous elasticity. By itself, this intertemporal effect could be due to reciprocity or momentum, as both predict higher effort following a higher piece rate.

Experiment 1 also has similar findings once restricted to the first three treatment periods, as can be seen in Table 4A. This may be due to the fact that in the first experiment, every individual received all 3 treatment pairs. As discussed in Section 4, this suggests the presence of multiple treatment interactions that were not ex ante predicted. Therefore, to limit the analysis to the post-treatment effects of just the first treatment pair, I analyze only the first three treatment periods.[59] Upon doing

model. But if the targets are not strong enough to induce behavior that differs, their predictive value is reduced.

[58] Contrary to this prediction, the literature has found some evidence of effects in surrounding periods in the pursuit of other research (Cardella and Depew (2015); Bradler et al. (2015); Connolly (2008)).

[59] Restricting analysis to only the first treatment pair is equivalent to focusing on the first 3 periods; however as half of those subjects received treatment in period 3, the following period would be period 4. Results change very little when limiting it to individuals who have only received baseline piece rate (0.05) and leisure (YouTube) in period 4.

so, estimates suggest that a 5 cent (100%) piece rate increase in the previous period increases effort by 0.5 - 0.66 correct problems (6 - 9%). By itself, this result could be indicative of either momentum or reciprocity, as shown in Section 2.

Also worth noting is that while cellphones do not effect effort on average, it does seem to reduce contemporaneous effort for men over both experiments (see Tables 5A and 5B). This effect also persists in experiment 2 and some specifications of experiment 1. Also worth noting is that while cellphone *access* does not significantly alter effort in future periods, that self-reported cellphone *usage* is correlated with decreased effort in future periods, even after controlling for worker ability with productivity proxies (see Appendix Table 6).[60] Therefore, the intertemporal evidence of leisure is broadly suggestive of momentum rather than reciprocity (or other theories), as reciprocity would suggest a worker work harder after use or access to an increased leisure technology, not less hard.[61]

These findings also reject the "total income" target model. In this model, an agent receives relatively greater disutility for falling short of a particular income over multiple periods (in this case, the experimental session). If subjects in this experiment exhibited a total income reference point, subjects should reduce effort following a high piece rate period, as the agent was more likely to have hit their target in the preceding period. As shown in Tables 4A and 4B, the previous piece rate is instead positively correlated with effort in this period. Thus, I conclude there was no significant evidence of a total income target in this experiment.[62]

Lastly, being informed of the upcoming piece rates one period in advance did not influence effort. In the results of experiment 1, presented in Table 6A, there seems to be some borderline significant results when focusing on early panels, but not in the full panel. In the results of experiment 2, presented in Table 6B, knowledge of future piece rates also has no effect on effort – despite having similar sized standard errors and a similar level of power to detect as the effect of past

[60] However, this finding has the potential for selection effects driving omitted variable bias, suggesting the coefficients should not be taken as causal estimates.

[61] As reciprocity seems more likely to trigger with non-monetary goods (Kube et al. (2012)), one might expect cell phone access to be even more likely to generate reciprocity than increased piece rates. One possible caveat – if subjects engage in cellphone use but do not actually "enjoy" this ability to use cell phones, e.g. due to self-control problems, it may not necessarily engage in reciprocity. However, this does not seem to be the case as the subject has a number of other self-control methods for cell phones (turning phone off, pulling out battery, leaving at home) and otherwise might suffer a very similar self-control issue with YouTube.

[62] Although not detected in my setting, Abeler et al. (2011) have devised an elegant method to elicit loss aversion even in a laboratory setting by varying a random outside option. It may be that in my experiment the subject has no experience with the task prior to the experiment, making it difficult to select a total income target.

piece rates. This evidence is suggestive of naive momentum rather than sophisticated momentum or reciprocity; both of these alternatives would predict effort increases upon learning about future piece rate increases.

Indeed, for both sophisticated momentum and reciprocity, one might expect the "pre-" piece rate effect to be larger than the "post-" piece rate effect. For reciprocity, work by Gneezy and List (2006) suggests that reciprocity decreases over time, thus the "post-" period having a larger effect is unlikely. For sophisticated momentum, extra effort in the "pre-" piece rate period would help take full advantage of the higher piece rates in the next period, whereas "post-" piece rate effects would be a result of previously expanded effort. Thus, the small and insignificant coefficient of future knowledge is strong evidence in favor of Naive Momentum.

1.5.3. Instrumental Variable Approach

If naive effort momentum is occurring, Proposition 3 in Section 2 guides how one might estimate it – in particular using an AR(1) approach. As discussed in Section 4, assume the true model is of the following sort:

$$e_{i,t} \quad = \quad \rho \cdot e_{i,t-1} + f_i + \beta_1 w_{i,t} + \beta_2 \gamma_{i,t} + \nu_{i,t}$$

Where $e_{i,t}$ is the number of problems solved by individual i at time t, ρ captures the degree of "momentum" from the previous period, f_i is the individual ability or motivation component, $w_{i,t}$ is the piece rate at time period t and $\gamma_{i,t}$ is the leisure technology available at time t, and $\nu_{i,t}$ is an error term. This theory allows us to encapsulate the force of momentum in a single parameter, which is potentially broader in application and policy implications and allows for easier comparisons across tasks (Charness and Kuhn (2011)).

While one could design an OLS structure to estimate the above, as mentioned before, the presence of fixed effects may bias the parameter ρ. I can employ productivity proxies as in previous specifications, but there may be remaining omitted variable bias as the uncaptured component of f_i , which now resides in the error term, may be correlated with $e_{i,t-1}$.

However, natural instrumental variables for previous period's effort are present – the previous period's piece rate and leisure technology. These variables are assigned randomly, but should influence the previous period's effort directly. To achieve asymptotic consistency, the instrumental variable

w_{it-1} would need to satisfy the following:

$$\begin{aligned} Cov(w_{it-1}, e_{it-1}) &\neq 0 \quad \text{("First stage")} \\ Cov(w_{it-1}, \nu_{it}) &= 0 \quad \text{("Exclusion Principle")} \end{aligned}$$

While the "first stage" is strong as piece rates do impact contemporaneous effort (see Tables 3A and 3B), one might have reasonable doubts about the exclusion principle. In particular, suppose the true data generating process was a model of reciprocity, a process in which past piece rates directly influence current effort (rather than influencing effort through past effort). For example,

$$e_{it} = \rho e_{it-1} + \alpha_1 w_{it} + \alpha_2 w_{it-1} + \alpha_3 \gamma_{it} + \nu_{it}$$

If data from this data generating process was used to estimate an AR(1) model without w_{it-1} as a regressor, then $\alpha_2 w_{it-1}$ would remain in the error term. Since w_{it-1} will be correlated with e_{it-1}, this would result in omitted variable bias. In this case, it would overestimate the magnitude of ρ, as what is actually driven by reciprocity would be misinterpreted as momentum (w_{it-1} and e_{it-1} positively correlated).

Therefore, in order to believe the asymptotic consistency of an instrumental variable (IV) approach, one must be reasonably confident that the other models where previous piece rates enters directly (such as reciprocity or "total" income targeting) are not occurring. Although momentum most closely fits the comparative statics, additional discussion of alternate theories is provided below.

With this caveat in mind, I apply the instrumental variables (IV) approach using previous piece rate and phone access to predict previous period's effort. As presented in Table 7, the estimates find around 43-45% of the increased effort is retained in the following period, even once incentives revert to baseline.[63] I replicate this estimate of 43% using an alternate slider task in a replication experiment (see Appendix Table 1 and Section 10.3). Experiment 1 suffers from a weak instrument problem due to a smaller sample, but several specifications of experiment 1 are in line with this estimate of 45% (see Appendix Table 2).

[63]This is substantially higher than the estimate of 75% given by OLS without employing an IV strategy.

1.6. Alternative Theories

While momentum seems to be the most parsimonious description of the contemporaneous and intertemporal results, careful consideration of other theories is warranted.

To reiterate, the comparative statics for the most part the design gives us have two tests to differentiate between momentum and reciprocity, both of which point in the direction of momentum.

1. As discussed in Section 5.2, informing a subject of a future piece rate increase or leisure option did not immediately increase effort as predicted by a model of reciprocity. Instead, I find that individuals only increase effort once the higher piece rate is applied. See Tables 6A and 6B.
2. After cell phone access, subjects exert less effort, not more as predicted by reciprocity (see Section 2.3 for predictions). See Tables 5A and 5B as well as Appendix Table 6 for results.

One possible alternative explanation for the experimental findings is that workers who experienced additional problems were able to increase their productivity in post-treatment periods ("On the Job Learning"). If there is significant on the job learning, then increased effort in an early period could result in additional problems solved in later periods. If true, this could account for stickiness detected.

However on-the-job learning seems an unlikely explanation as there was no indication of increased productivity over time as one would predict – see figures 2A and 2B. In addition, one might expect the number of incorrect problems to fall over time with learning, but this does not happen. Also, the task itself (counting 100 images dozens of times) has a limited scope for learning – indeed, Figure 4 shows a rapid convergence of problems solved per minute even within the Pre-treatment period.

Furthermore, even if on-the-job learning were occurring, it seems unlikely to explain the post-treatment effects. The relatively small increase in problems solved during period 2 (approximately 1 problem) is small relative to the total number of problems solved by that point (approximately 3%). If increasing the number of problems solved by 3% increases productivity by 10%, then period 3 should see an increase of approximately 60% even in the control group, which is inconsistent with the data. In addition, if this on the job learning was known to workers, then a future increase in piece rate would motivate additional effort in the preceding period so as to increase productivity. This was also not found in the data.

To recap why a period income targeting model does not fit the data, one would expect an increased piece rate to decrease contemporaneous effort (if the period income target is a significant component of utility). In addition, without adaptive references, a period income target model would predict no intertemporal spillovers. For more details, see the discussion in the contemporaneous effects section above.

To address why a "total" income targeting model does not fit the data, note that an increase in piece rate should reduce effort in surrounding periods. A higher piece rate makes it easier to hit a fixed "total" income target. Thus, to the extent that income targets induce effort,[64] the worker would exert less effort in the lower piece rate periods compared to control.[65] Instead, the data displays an increase in effort. For more details, see the discussion in the intertemporal results section above.

However, another possibility is period-level income reference dependence with adaptive references. In other words, by earning more in the previous period, the agent increases the income target for the following period.[66] This model may be hard to distinguish empirically from momentum, but there are two related tests that suggest adaptive income references are not driving the results.

First, if references are an important component of the utility function, this model would predict decreased effort when faced with a higher piece rate. This occurs as it is now easier to hit the income reference of the previous period. In the data, there are no such decreases when faced with higher piece rates, and workers work less hard when piece rates decrease.

Second, every subject had a counter to keep track of earnings from that period (see Figure 1). This was implemented to reduce subject uncertainty about earnings. In addition, experiment 2 subjects either had a "previous period earnings" or a "total earnings" counter located below the period earnings. This was randomly assigned at an individual level to potentially nudge period or daily income targeting. Specifically if an individual had been given the "previous period earnings" counter treatment and were driven by a period-level effort reference model, the post-treatment effects should have been stronger as they have more precise information about the effort and earnings exerted in

[64] If the target is too low or too high (e.g. $0.10 or $1000 for a laboratory study), then the agent will demonstrate behavior consistent with a neoclassical time separable model as the kink in utility will not be relevant to effort decisions.

[65] A neoclassical time separable model with income effects has similar predictions.

[66] Although not discussed extensively in this paper, a theory of adaptive income references presented in Brandon et al. (2014); Kőszegi and Rabin (2006, 2007, 2009) would also generally have an effect if information about piece rates is presented in advance. In addition, as shown by Brandon et al. (2014); Huffman and Goette (2006) workers who receive a higher lump sum early in the day should reduce their optimal effort afterward. This does not fit with the findings above, as workers treated to a higher piece rate worked harder even after incentives return to baseline.

the previous period. As can be seen in Table 8, this information did not significantly change the impact of an increased piece rate, either contemporaneously or in the previous period.[67] Though it seems to have influenced phone use slightly, this is hard to construe as evidence consistent with income reference dependence.

Another possibility is worker confusion regarding the piece rates. There are two reasons why this is unlikely to be driving results. First, before every period, the worker is presented a new instructions page which clearly outlines the piece rate in that period. This instructions page cannot be skipped for at least 30 seconds and workers must successfully type in the piece rate before they can continue. If the worker has information about future incentives, they are also quizzed on the future piece rate. Second, as mentioned above, in all experiments there was a counter that showed how much the subject had earned that period; thus even if they failed to understand the instructions, subjects would quickly see how much each problem was earning them. Of the 422 workers in the second experiment, only 4 individuals answered that the compensation was "somewhat unclear" or "unclear" in a post-experiment survey.

One last potential explanation is a lack of worker trust. If workers do not trust the promised piece rate increase, they may not reciprocate it initially, and instead wait until the piece rate is actually put in place. However, I believe two aspects make this explanation unlikely. First, 90% of subjects have previously completed 3 or more studies at the Wharton Behavioral Lab. As a dedicated experimental lab, Wharton Behavioral Lab has a reputation and incentives for upholding its promises to subjects. Secondly, if the worker did not trust promises of higher piece rates, it is unclear why higher piece rates would incentivize them to work harder in the treated period either, as there was no actual payment until the end of the experimental session.

1.7. Conclusion

I investigated the intertemporal elasticity of labor supply with a series of incentivized real effort experiments and find effort levels persist even once incentives return to baseline. After testing predictions to distinguish theories, I find strong evidence of effort momentum over short time scales and estimate a 5-minute momentum parameter of 0.45 across multiple experiments and tasks. This suggests it takes 15 minutes after an interruption to return to 90% of prior productivity levels, in line with observational evidence on interrupted work (Mark et al. (2005)). Providing information

[67] Unfortunately, while every subject did face a randomized counter, a small programming typo prevented the capture of this variable for the first day. As it is unclear which counter day 1 subjects faced, they are dropped from Table 7.

a full period in advance of does not seem to significantly influence this effort allocation – further suggesting a "naive" sort of momentum.

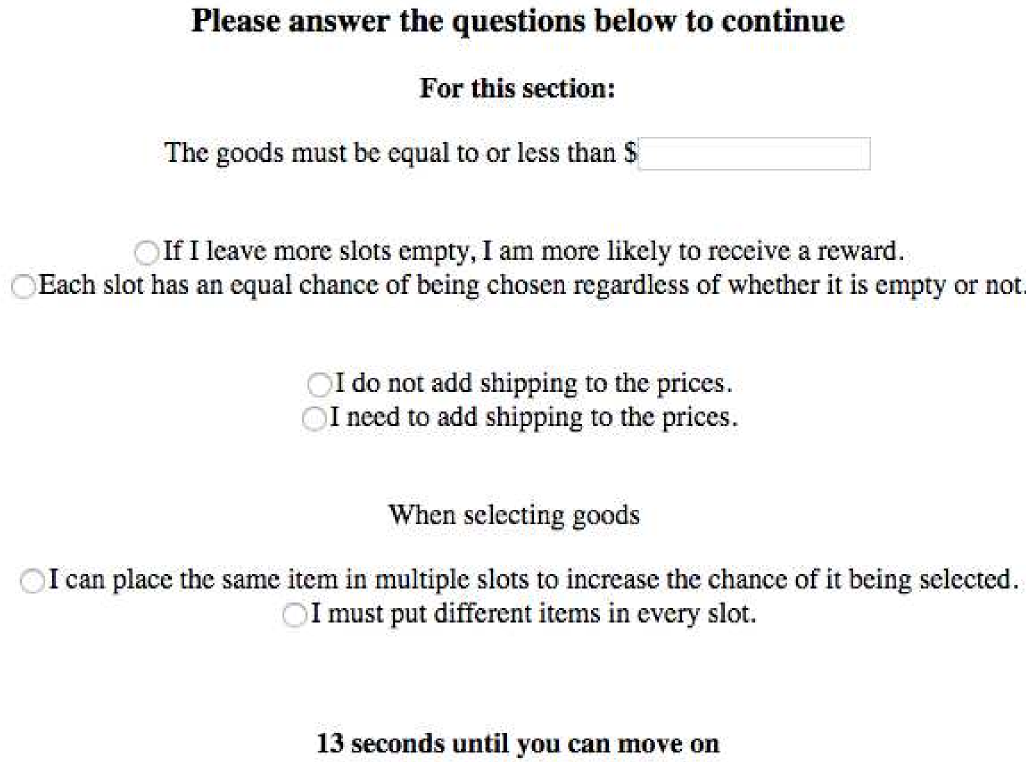
One weakness of this study is remaining uncertainty regarding the source of the momentum effects. For example, if effort momentum is a result of quickly decaying task-specific human capital (i.e. a "train of thought"), then switching tasks could be equally harmful as being interrupted. This would also be consistent with evidence that multitasking is less productive than sequential work, as found in Buser and Peter (2012). Alternatively, it may be that momentum has a physiological component, perhaps due to adrenaline or other neurobiological processes. Lastly, the momentum could also be related to effort reference dependence (as opposed to income reference dependence), though many reference dependence models would predict that receiving information in advance should change the effort allocation (Brandon et al. (2014); Kőszegi and Rabin (2006, 2007, 2009)). Distinguishing these theories could help provide additional suggestions on how to minimize momentum loss after an interruption, e.g. a cellphone wallpaper reminding one to return to work after a phone call or doing 5 jumping jacks immediately after an interruption.

There's also some uncertainty to the extent to which workers are aware of these momentum effects. Although they do not seem to employ information to take advantage of momentum, there is still a chance workers are aware of it conceptually. As has occurred with some past studies (Price and Wolfers (2010); Pope et al. (2013)), increased awareness of the momentum effect may overturn or undo some of the effect. For example, if a worker knows they tend to work harder after working hard, they may slack off early and expect the work to "finish itself". Or as Mark et al. (2008) find, workers may work harder following an interruption to "catch up", though I find no evidence of this. One possibility to investigate the degree of self-awareness is to use costly commitment with a self-selected cut-off, akin to Kaur et al. (2010).

Another open question is whether these momentum effects would persist over longer time periods. One replication experiment with multiple periods following the treatment suggests that effort continues to decay exponentially, suggesting that the effects of momentum would disappear within 20 minutes or so. That being said, even if short-lived, measuring momentum may have direct applications to the economics of task juggling and interruptions. As outlined in Section 5.3, approximately 45% of effort momentum persists after 5 minutes. Using this estimate as a starting point, 45% of productivity is lost in the first 5 minutes after an interruption, an additional 20% in the second 5 minutes, 9% in the third 5 minutes, 4% in the next 5 minutes and so on. In total, if an interruption causes me to lose 5 minutes of productivity, I lose an additional 4 minutes of productivity due to effort momentum loss spread out over the next 30 minutes. Put in other terms, total productivity loss from effort momentum is 80% of the original interruption loss. Given estimates of the number of interruptions knowledge workers face, that suggests up to an hour of productivity per work day could be lost due to effort momentum alone.

1.8. Figures

Figure 1: Example of Counting Problem Task



Please answer the questions below to continue

For this section:

The goods must be equal to or less than $

If I leave more slots empty, I am more likely to receive a reward.
Each slot has an equal chance of being chosen regardless of whether it is empty or not.

I do not add shipping to the prices.
I need to add shipping to the prices.

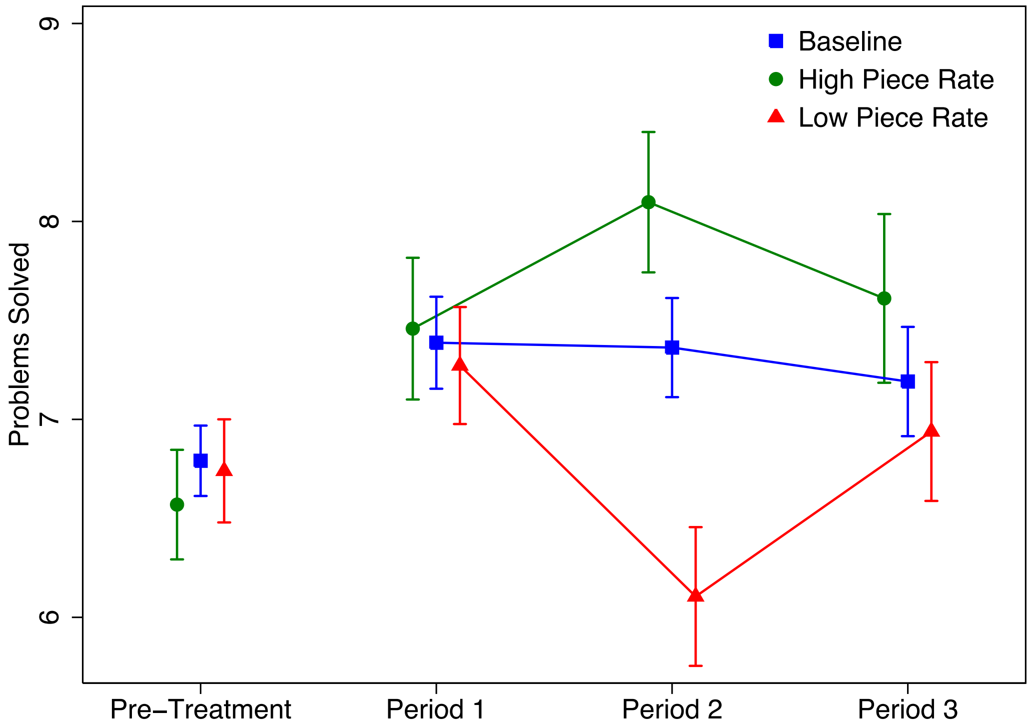
When selecting goods

I can place the same item in multiple slots to increase the chance of it being selected.
I must put different items in every slot.

13 seconds until you can move on

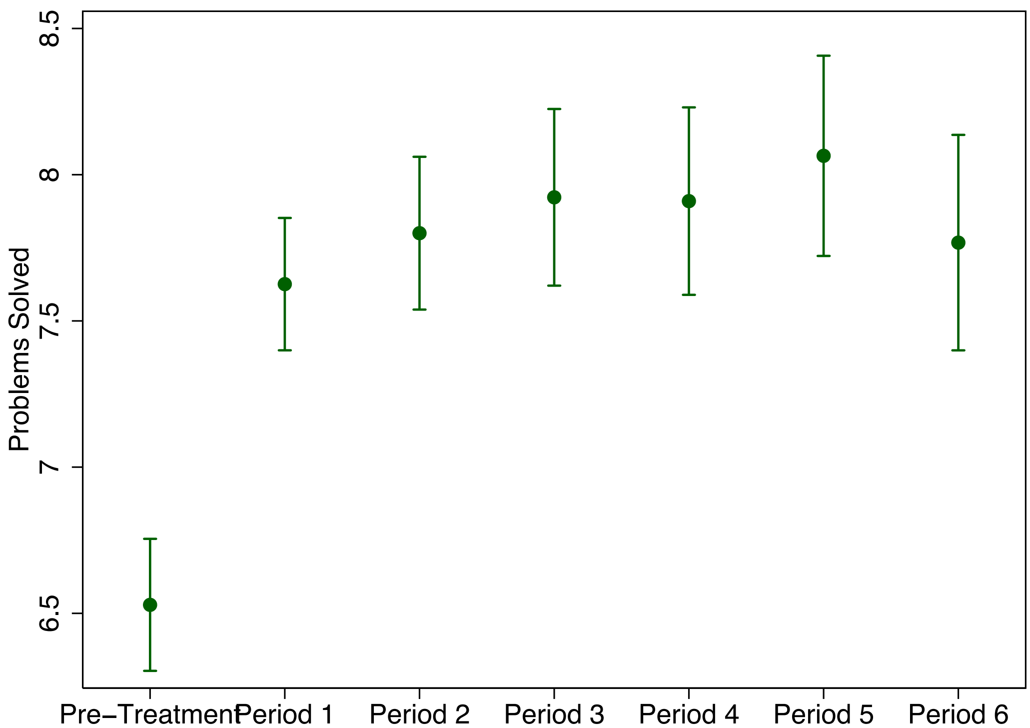
Notes. Figure demonstrates a typical counting task screen faced by subject. Whether subject was asked to count hearts or drops was randomized in experiment 2.

Figure 2: Solved Problems by Period – Experiment 2



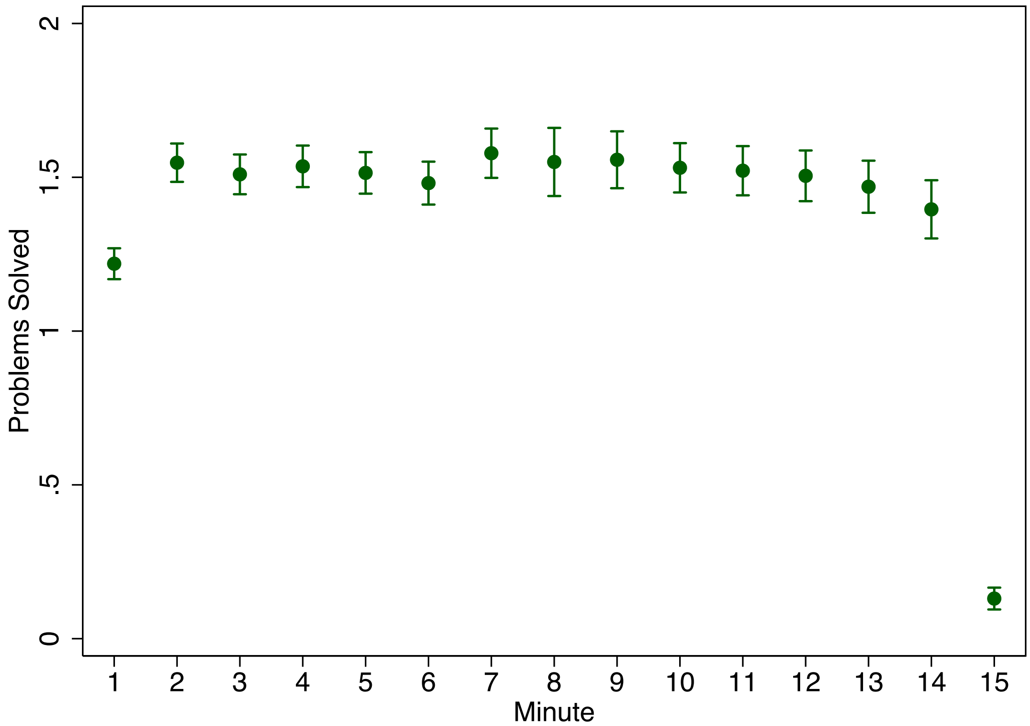
Notes. Bars represent standard errors. Vertical axis represents the number of problems solved by workers. Treatments were only in effect for period 2 (see experimental design in section 3). The Pre-Treatment period is a training period to familiarize workers with the task. Pre-Treatment lasted 3 times the duration of the other periods and thus the problems solved in Pre-Treatment is divided by 3 to provide accurate comparison.

Figure 3: Solved Problems by Period – Experiment 1



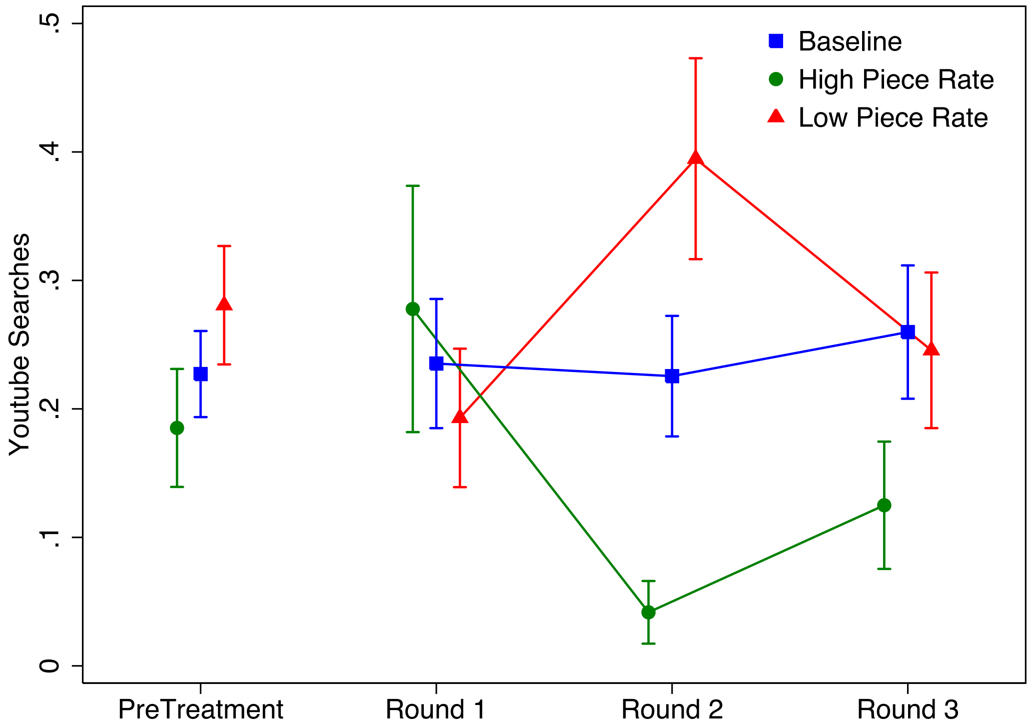
Notes. Bars represent standard errors. Three treatment pairs were applied at varying periods (see experimental design in section 3). The Pre-Treatment period is a training period to familiarize workers with the task.

Figure 4: Solved Problems within Pre-Treatment Period – Experiment 2



Notes. Bars represent standard errors. This figure demonstrates the number of problems solved by minute of the pre-treatment period. As these counting problems take about 45 seconds, the final minute was lower due a mechanical effect (of being unable to finish a problem in time) and additional uncertainty of whether one is able to finish the problem in time (perhaps due to the timer reading “0 minutes left”).

Figure 5: YouTube Searches by Period – Experiment 2



Notes. Bars represent standard errors. Vertical axis represents the number of YouTube searches performed by workers (baseline leisure option). Treatments were only in effect for period 2 (see experimental design in section 3). The Pre-Treatment period is a training period to familiarize workers with the task. Pre-Treatment lasted 3 times the duration of the other periods and thus the problems solved in Pre-Treatment is divided by 3 to provide accurate comparison.

1.9. Tables

Table 3: Summary Statistics

	Experiment 1				Experiment 2			
	Mean	Standard dev	Min	Max	Mean	Standard dev	Min	Max
Individual Level Variables								
Female	0.72	0.45	0	1	0.71	0.45	0	1
Problems Solved in PreTreatment	6.59	2.75	0	14	20.2	7.60	0	67
Age	21.3	5.27	18	61	20.4	1.85	18	38
SAT Math Score	731	78	165	800	731	63	400	800
Total Payment	3.87	1.75	0	10.75	2.16	0.94	0	5.4
Computer Skill Test	2.01	0.08	2	3	2.01	0.10	2	3
Number of Previous Lab Studies	33.4	26.7	0	129	23.7	25.0	0	292
Period Level Variables								
Problems Solved	7.85	3.7	0	21	7.23	3.49	0	17
Problems Incorrect	0.06	0.29	0	4	0.08	0.30	0	3
Youtube Searches	0.16	0.59	0	6	0.24	0.67	0	5
Period Payment	0.60	0.65	−0.65	5.4	0.40	0.30	−0.1	2.1
High Piece Rate Indicator	0.17	0.37	0	1	0.08	0.28	0	1
Phone Access Indicator	0.17	0.37	0	1	0.09	0.29	0	1
Low Piece Rate Indicator	n.a.	n.a.	n.a.	n.a.	0.08	0.27	0	1
Number of Individuals	155				422			
Number of Treatment Periods	930				1266			

Notes. Computer Skill Test was a demographic variable collected by the Wharton Behavioral Lab prior to the experiment, however one with almost no variation. SAT Math score is missing for individuals who either took the ACT or otherwise did not wish to share that information with researchers. Indicators for treatments are presented under the period level variables – as experiment 1 had no "low piece rate" treatment, it has no such indicator.

Table 4: Randomization Check – Experiment 1

Dependent Variable	Period # for			
	Piece Rate Treatment		Phone Treatment	
Female	-0.09	0.19	0.06	0.03
	(0.29)	(0.32)	(0.30)	(0.32)
SAT Math Score		-0.002		-0.001
('00s of points)		(0.002)		(0.26)
PreTreatment Problems Solved		-0.047		0.011
		(0.058)		(0.048)
F-test	0.10	1.24	0.05	0.15
p value	0.75	0.30	0.83	0.93
Dependent Variable Mean	3.34	3.41	3.48	3.54
Number of Observations	930	738	930	738
Number of Individuals	155	123	155	123
Adj-R^2	0.001	0.018	0.001	0.005

Notes. Standard Errors (clustered at individual level) presented in parentheses above. As every subject in experiment 1 receives all treatments at some point, the dependent variable is the period in which they received the treatment in question. If randomization was done properly, the pre-treatment variables should not predict the period they received the treatment. Indeed, the F-stats are all large enough that I fail to reject the hypothesis that all coefficients are zero under $\alpha = 0.05$. Thus, I conclude the randomization was adequately done. SAT Math score is missing for 32 individuals who either took the ACT or otherwise did not wish to share that information with researchers.

Table 5: Randomization Check – Experiment 2

Variable	Baseline	Piece Rate Decrease	Piece Rate Increase	Phone Access	
Female	0.68	0.71	0.74	0.71	$p < 0.76$
	(0.47)	(0.46)	(0.44)	(0.45)	(F-test = 0.39)
Age	20.26	20.13	20.73	20.39	$p < 0.11$
	(1.75)	(1.61)	(2.44)	(1.48)	(F-test = 2.05)
# Previous Studies at Lab	25.18	24.1	26.23	23.64	$p < 0.22$
	(23.9)	(20.5)	(34.5)	(18.7)	(F-test = 0.88)
Computer Skill Test	2.01	2.01	2.02	2.00	$p < 0.56$
	(0.01)	(0.09)	(0.14)	(no variation)	(F-test = 0.56)
Problems Solved in PreTreatment	20.57	20.23	19.54	20.23	$p < 0.80$
	(7.76)	(8.2)	(7.02)	(7.22)	(F-test = 0.33)
Number of Subjects Treated	103	114	104	101	

Notes. As every subject in experiment 2 receives (at most) one primary treatment, the subjects are split according to primary treatment. Means and standard deviations (in parentheses) are presented by primary treatment. If randomization was done properly, the pre-treatment variables should not differ significantly according to which treatment was received. Indeed, for all rows the F-stat corresponds to a p greater than 0.05 (fail to reject the hypothesis that all coefficients are less than zero under $\alpha = 0.05$).

Table 6: Contemporaneous Piece Rate and Phone Access: Impact on Effort – Experiment 1

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	4.34***	4.62***	4.62***	2.07*	2.07*
(in cents per problem)	(1.17)	(1.15)	(1.14)	(1.09)	(1.09)
Phone Access	−0.38**	−0.37**	−0.39**	−0.46**	−0.46**
	(0.19)	(0.19)	(0.19)	(0.19)	(0.19)
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.85	7.85	7.85	7.85	7.85
Number of Observations	930	930	930	930	930
Number of Individuals	155	155	155	155	155
Adj-R^2	0.01	0.23	0.23	0.31	0.32

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include sex, age, ethnicity bins, number of sessions done, and WBL computer diagnostic scores. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 7: Contemporaneous Piece Rate and Phone Access: Impact on Effort – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	12.5***	14.9***	16.6***	16.7***	16.7***
(in cents per problem)	(3.39)	(2.67)	(3.02)	(2.98)	(2.98)
Phone Access	−0.05	−0.10	0.12	0.15	0.17
	(0.33)	(0.25)	(0.32)	(0.32)	(0.32)
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.23	7.23	7.23	7.23	7.23
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.01	0.41	0.41	0.43	0.45

Notes. The dependent variable is the number of problems solved correctly in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include sex, age, ethnicity bins, number of sessions done, and WBL computer diagnostic scores, but could not be matched for 2 subjects. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 8: Previous Period Piece Rate and Phone Access: Impact on Effort – Experiment 1

Dependent Variable:	Specification					
Problems Solved	(1)	(2)	(3)	(4)	(5)	(6)
Piece Rate	5.00***	5.34***	2.19*	8.06***	7.74***	5.43***
(cents per problem)	(1.29)	(1.24)	(1.14)	(2.48)	(1.94)	(1.87)
Previous Period's Piece Rate	3.42*	3.86**	0.45	7.99**	8.04***	5.91**
(cents per problem)	(1.74)	(1.56)	(1.36)	(3.59)	(2.93)	(2.78)
Phone Access	−0.26	−0.25	−0.43**	−0.29	−0.30	−0.38
	(0.21)	(0.21)	(0.20)	(0.36)	(0.32)	(0.30)
Previous Period Phone Access	0.24	0.28	0.07	0.22	0.18	0.21
	(0.26)	(0.24)	(0.23)	(0.41)	(0.35)	(0.34)
PreTreatment Quintiles		X	X		X	X
Period Fixed Effects		X	X		X	X
Session Fixed Effects			X			X
Individual Controls			X			X
Periods 1 to 3 Only				X	X	X
Dependent Variable Mean	7.85	7.85	7.85	7.79	7.79	7.79
Number of Observations	930	930	930	465	465	465
Number of Individuals	155	155	155	155	155	155
Adj-R^2	0.01	0.24	0.32	0.04	0.34	0.39

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. "Periods 1 to 3" uses data of the first treatment period and following period to minimize treatment interactions. Individual Controls include sex, age, ethnicity bins, number of sessions done, and WBL computer diagnostic scores. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 9: Previous Period Piece Rate and Phone Access: Impact on Effort – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	12.66***	15.09***	16.57***	16.97***	17.09***
(cents per problem)	(3.51)	(2.74)	(3.03)	(3.05)	(3.11)
Previous Period's Piece Rate	4.09	6.51**	7.24**	7.64**	7.87**
(cents per problem)	(4.03)	(3.11)	(3.32)	(3.35)	(3.40)
Phone Access	−0.03	−0.08	0.12	0.16	0.18
	(0.36)	(0.28)	(0.32)	(0.32)	(0.33)
Previous Period Phone Access	−0.03	−0.08	0.02	0.06	0.07
	(0.41)	(0.33)	(0.36)	(0.36)	(0.36)
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.23	7.23	7.23	7.23	7.23
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.01	0.41	0.41	0.44	0.45

Notes. The dependent variable is the number of problems solved correctly in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 2 subjects. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 10: Phone Access by Gender: Impact on Effort – Experiment 2

Dependent Variable:	Specification					
Problems Solved	(1)	(2)	(3)	(4)	(5)	(6)
Phone Access * Female	−0.26	−0.28	−0.29	0.08	−0.18	−0.02
	(0.22)	(0.22)	(0.22)	(0.41)	(0.35)	(0.32)
Previous Period Phone * Female	0.35	0.32	0.29	0.12	−0.39	−0.08
	(0.25)	(0.24)	(0.24)	(0.45)	(0.40)	(0.38)
Phone Access * Male	-1.10^{***}	-1.10^{***}	-1.11^{***}	-2.33^{***}	-1.78^{**}	-2.02^{***}
	(0.39)	(0.41)	(0.41)	(0.77)	(0.73)	(0.72)
Previous Period Phone * Male	−0.69	−0.59	−0.65	−0.42	0.42	0.18
	(0.57)	(0.55)	(0.55)	(1.05)	(0.69)	(0.71)
Male	-1.02^{*}	−0.68	−0.10	−0.36	−0.28	0.24
	(0.58)	(0.58)	(0.56)	(0.61)	(0.50)	(0.54)
Pre-Treatment Quintiles		X	X		X	X
Period Fixed Effects		X	X		X	X
Session Fixed Effects			X			X
Individual Controls			X			X
Periods 1 to 3 Only				X	X	X
Dependent Variable Mean	7.85	7.85	7.85	7.79	7.79	7.79
Number of Observations	930	930	930	465	465	465
Number of Individuals	155	155	155	155	155	155
Adj-R^2	0.03	0.24	0.33	0.03	0.32	0.39

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. "Periods 1 to 3" uses data of the first treatment period and following period to minimize treatment interactions. Individual Controls include sex, age, ethnicity bins, number of sessions done, and WBL computer diagnostic scores. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 11: Phone Access by Gender: Impact on Effort – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Phone Access * Female	0.51	0.36	0.39	0.44	0.45
	(0.36)	(0.29)	(0.32)	(0.34)	(0.35)
Previous Period Phone * Female	0.48	0.34	0.42	0.47	0.48
	(0.44)	(0.36)	(0.39)	(0.40)	(0.40)
Phone Access * Male	−1.68**	−1.59***	−1.57***	−1.61***	−1.64***
	(0.82)	(0.51)	(0.54)	(0.51)	(0.53)
Previous Period Phone * Male	−1.61*	−1.52**	−1.43**	−1.48**	−1.51**
	(0.88)	(0.61)	(0.63)	(0.61)	(0.61)
Male	−0.45	−0.48	−0.48	−0.51*	−0.53*
	(0.36)	(0.27)	(0.27)	(0.28)	(0.28)
Pre-Treatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.23	7.23	7.23	7.23	7.23
Number of Observations	1263	1263	1263	1263	1260
Number of Individuals	421	421	421	421	420
Adj-R^2	0.02	0.42	0.42	0.44	0.44

Notes. The dependent variable is the number of problems solved correctly in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, ethnicity, computer skill test, and total # of experimental sessions done at the lab. Gender could not be matched for one subject, and the controls for an additional subject. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 12: Next Period Piece Rate and Phone Access: Impact on Effort – Experiment 1

Dependent Variable:	Specification					
Problems Solved	(1)	(2)	(3)	(4)	(5)	(6)
Piece Rate	5.33***	5.75***	2.78**	8.74***	8.62***	6.46***
(in cents)	(1.26)	(1.23)	(1.20)	(2.58)	(1.98)	(1.95)
Next Period Piece Rate	1.80	2.65	0.57	3.63*	5.01**	4.20*
(if known)	(1.73)	(2.23)	(2.35)	(2.01)	(2.14)	(2.47)
Previous Period Piece Rate	3.64**	3.90***	0.84	8.75**	8.49***	6.53**
	(1.62)	(1.45)	(1.33)	(3.73)	(2.97)	(1.62)
Pre-Treatment Quintiles		X	X		X	X
Period Fixed Effects		X	X		X	X
Shown Next Period Bin		X	X		X	X
Session Fixed Effects			X			X
Individual Controls			X			X
Periods 1 to 3 Only				X	X	X
Dependent Variable Mean	7.85	7.85	7.85	7.79	7.79	7.79
Number of Observations	930	930	930	465	465	465
Number of Individuals	155	155	155	155	155	155
Adj-R^2	0.01	0.23	0.32	0.04	0.34	0.39

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. "Only Periods 1 to 3" uses data of the first treatment period and following period to minimize treatment interactions. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 13: Next Period Piece Rate and Phone Access: Impact on Effort – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	12.8***	15.5***	16.3***	16.65***	16.76***
(in cents)	(3.58)	(2.79)	(2.92)	(2.96)	(3.04)
Next Period Piece Rate	1.09	3.23	0.73	1.59	2.63
(if known)	(2.64)	(2.19)	(2.62)	(3.92)	(3.88)
Previous Piece Rate	4.21	6.88**	7.20**	7.52**	7.75**
	(4.12)	(3.17)	(3.27)	(3.31)	(3.38)
Pre-Treatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Shown Next Period Bin				X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.23	7.23	7.23	7.23	7.23
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.01	0.41	0.41	0.44	0.44

Notes. The dependent variable is the number of problems solved correctly in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 2 subjects. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 14: Previous Effort Instrumental Variable: Impact on Effort – Experiment 2

Dependent Variable:			Specification		
Problems Solved	(1)	(2)	(3)	(4)	(5)
Problems Previous Period	0.39	0.50***	0.45***	0.44***	0.43***
	(0.26)	(0.18)	(0.17)	(0.17)	(0.17)
Piece Rate	12.7***	13.9***	15.3***	15.2***	14.9***
	(2.63)	(2.51)	(2.73)	(2.81)	(2.88)
Phone Access	0.04	0.07	0.24	0.25	0.23
	(0.24)	(0.23)	(0.27)	(0.27)	(0.27)
First Stage F Stat (IV)	6.1	14.5	15.5	16.5	16.2
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.22	7.22	7.22	7.22	7.22
Number of Observations	844	844	844	844	840
Number of Individuals	422	422	422	422	420
Adj-R^2	0.39	0.56	0.57	0.57	0.57

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from linear Instrumental Variable regressions estimated by (iterative) GMM and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 2 subjects. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 15: Period or Total Earnings Salience: Impact on Earnings – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	9.32*	13.12***	17.05***	15.84***	15.68***
	(5.27)	(3.83)	(4.17)	(3.87)	(3.88)
Piece Rate * Period Salience	2.41	−3.82	-2.86	−1.01	−1.00
	(6.93)	(5.21)	(5.33)	(5.11)	(5.13)
Phone Access	0.92*	0.62*	0.94**	0.92**	0.88**
	(0.50)	(0.35)	(0.41)	(0.46)	(0.46)
Phone Access * Period Salience	−2.03***	−1.20**	−1.21**	−1.23**	−1.15**
	(0.73)	(0.53)	(0.54)	(0.56)	(0.54)
Period Salience	-0.32	0.34	0.40	0.22	0.25
	(0.58)	(0.40)	(0.39)	(0.39)	(0.39)
Pre-Treatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.28	7.28	7.28	7.28	7.28
Number of Observations	894	894	894	894	891
Number of Individuals	298	298	298	298	297
Adj-R^2	0.02	0.45	0.48	0.51	0.52

Notes. The dependent variable is the total earnings from a single period. All specifications report results from OLS regressions and also include a constant term. The subject is shown either the previous period's earnings (as indicated by "Period Salience") or shown total earnings up to that period. Experiment 2 was the only one that featured this variation. Unfortunately, while every subject in Experiment 2 did face a randomized period or total counter, a small programming typo prevented the capture of this variable for the first day. As it is unclear which counter subjects faced on the first day, they are dropped from analysis above. $* = p < 0.1$, $** = p < 0.05$, $* * * = p < 0.01$.

CHAPTER 2 : Risk Over Goods

2.1. Introduction

Many important decisions involve risk, including insurance, portfolio choice, and labor relations. As a result, researchers have made great strides in understanding how decision makers perceive and value these risks. Yet the standard decision making models generally rely on a (concave) utility function over a single wealth variable. This often allows us to encapsulate risk preferences in a single parameter, which allows for comparisons across contexts and individuals.

While parsimonious, these models of risk aversion simplify a great deal of decision making. Recent evidence suggests that estimates of risk aversion may not be applicable across all domains. In Einav et al. (2012), demand for different types of insurance appears to be correlated, but does not correlate well with riskiness of the 401(k) investments.[68] In Barseghyan et al. (2011), demand for insurance (as measured by deductibles) were substantially different over two different goods, houses and cars.

One potential explanation for these differences in risk preferences is that individuals might treat pure monetary uncertainty (e.g. 401k or stocks) fundamentally different than good uncertainty (e.g. insurance for cars or houses). This paper sets out to test precisely this implication through an experiment with a real world market place, Amazon.com.[69] Subjects choose either Amazon.com credit amounts ($5, $10, etc.) or Amazon.com goods (books, clothing, etc.) that total up to either $20 or $100. They allocate these credits or goods across uncertain states that have equal probabilities of occurring. To remove temporal concerns, all credit awarded is spent immediately after the uncertainty is resolved.[70]

[68] It may be worth noting that Einav et al. (2012) focuses on how individuals rank relative to their peers. As a result, they are less focused on testing "absolute" differences between risk preferences over different goods and instead at the reliability of how individual will rank relative to others.

[69] A wide variety of goods is available on Amazon.com, making this an ideal environment to study evidence of real risk preferences over goods.

[70] Subjects were quizzed on this and several other topics to ensure understanding. See design section for more details.

If individuals treat self-selected goods and time-allocated money identically without uncertainty, then with relatively weak assumptions, there should be no difference between the allocated distributions (with uncertainty). As a simple example, if given \$20 of credit and an individual allocates \$10 to each of the two states, they will receive \$10 of credit for sure. This \$10 could be used to purchase anything on Amazon.com under or up to \$10. Therefore, it might be suprising that when asked to choose good(s) whose prices are at most \$20, individuals often no longer choose two goods under \$10, but may instead choose a \$15 good and a \$5 good. If this was indeed the optimal allocation of goods, it may seem strange that the subject did not choose a \$15 credit and \$5 credit allocation instead. The theory explains why

Contrary to this prediction, subjects exhibited considerably more risk aversion when selecting credit. Subjects were four times as likely to place "equal" quantities with credit than they were with goods. Furthermore, the mean standard deviation of credit allocations was about two-thirds that of the mean standard deviation of good prices. To analyze whether these differences could be driven by price uncertainty, subjects are randomly forced to spend more time on Amazon.com, but this does not seem to influence the allocations (with a rather precise zero effect).

Although this is the first research to explicitly test this uncertainty equivalence, an earlier theoretical literature uncovered several implicit assumptions about uncertainty over goods and money. Grant et al. (1992) tackles this by assuming preferences over monetary lotteries are induced by underlying preferences over goods lotteries.[71] The paper then goes on to establish implications for what risk aversion over monetary lotteries implies about risk aversion over good lotteries. However, rather than assuming preferences over monetary lotteries are induced by underlying preferences for good lotteries, I outline precisely what assumptions will generate indifference between a monetary lottery and equivalent value (self-selected) good lotteries.

This study is not unique in its interest in how consumers may treat money and goods

[71] To the credit of Grant et al. (1992), they acknowledge alternate approaches in footnote 7, even though it was not the main focus of that study.

differently. An extensive literature on the endowment effect indicates that subjects, after receiving a good, value that good more than subjects who do not (c.f. Knetsch (1989); Kahneman et al. (1991); Bordalo et al. (2012)). There is also a growing literature on salience and its impact on utiltiy over goods. While research has recently explored the potential for salience in monetary lotteries as in Bordalo et al. (2010) or for goods under certainty (Bordalo et al. (2012); Kőszegi and Szeidl (2013); Gabaix (2014)). This study therefore might contribute important empirical evidence on how individuals aggregate preferences over salient goods to create a limited-rationality utility framework over risk.

Consumers also face decisions daily about whether to purchase products running promotional contests (Dhar and Simonson (1992)). These contests pose somewhat of a mystery, given that they often feature prizes rather than equivalent cash values. In practice, these prizes may be sold at reduced costs to the promoter, but this study also indicates another possibility – individuals may wish to engage in risk over goods but prefer to avoid risk with equivalent cash prizes.This may have important implications for government run lotteries, which often serve to fund public programs. By adding physical items to these lotteries, it may be possible to encourage risk seeking behavior from participants and generate additional revenues for publicly funded programs.

The remainder of the paper is organized as follows. Section 2 demonstrates theoretical predictions. Section 3 outlines the experiment design. Section 4 presents the results and Section 5 concludes.

2.2. Theory

In this section, I demonstrate that under perfect information of goods available and prices, risk preferences across money and goods should be the same in a static model.

Each state $s = 1, 2, ..., S$ occurs with a probability γ_s. There are goods $n = 1, 2, ...N$ which can be consumed in each state, $g_{n,s}$ an element of the good-specific set $G_n \subset \mathbb{R}_+$, as well as a monetary good for each state, $m_s \in \mathbb{R}_+$. Thus, any particular lottery L is defined by the vector $(\gamma_1, m_1, g_{1,1}, g_{2,1}, ..., g_{N,1}; \gamma_2, m_2, g_{1,2}, g_{2,2}, ..., g_{N,2}; ...; \gamma_S, m_S, g_{1,S}, g_{2,S}, ..., g_{N,S})$, an element of $[0,1] \times \mathbb{R}_+ \times G_1 \times G_2 \times ... \times G_N \times [0,1] \times ... \times G_N$. For simplicity, I will call this vector space $\mathbb{L}_\mathbb{S}$ where S refers to the set of states.

This can also be written as the combination of degenerate lotteries L_s, where $L_s \equiv (1, m_s, g_{1,s}, g_{2,s}, ..., g_{N,s}) \in \mathbb{L}_1$. Thus, for any lottery L, for shorthand we may write it as $L = (\gamma_1 L_1,\ \gamma_2 L_2,\ ...,\ \gamma_S L_S)$ where $\gamma_s L_s$ refers to $(\gamma_s, m_s, g_{1,s}, g_{2,s}, ..., g_{N,s})$. To make assumptions of state independence more plausible,[72] I also assume that $\sum \gamma_s = 1$.

In addition, for any given state s define the market state as a vector of prices $P_s = (p_{1,s}, p_{2,s}, ..., p_{n,s})$. The market consists of a vector that consists of the individual market states $P = (P_1, P_2, ...P_S)$.[73] The agent has preferences relation $\succsim_P$ over lotteries $\mathbb{L}_\mathbb{S}$ for a given market P.[74] For notational simplicity, if there are lotteries $A, B \in \mathbb{L}_R$ with $R < S$, I write $A \succsim_P B$ as a shorthand for $(A, \vec{0}) \succsim_P (B, \vec{0})$ where $\vec{0} \in \mathbb{L}_{S-R}$. In words, even though preferences are over the entire S states, I pad out the remaining states with zeroes to use the same preference relation.

In addition to the basic relation assumptions, I assume the preferences have two additional properties: (a) Monetary Equivalence Under Certainty and (b) Independence.

Monetary Equivalence Under Certainty.

(i) For degenerate lottery $L_s = (1, m_s, g_{1,s}, g_{2,s}, ..., g_{N,s})$, the agent weakly prefers the bundle $L_s^{'} = (1,\ m_s + \sum_n p_{n,s} g_{n,s}, 0, 0, ..., 0)$, that is $L_s^{'} \succsim_P L_s$.

(ii) For any degenerate lottery $L_s = (1, m_s, g_{1,s}, g_{2,s}, ..., g_{N,s})$, there exists a degenerate lottery $L_s^{''} = (1, 0, g_{1,s}^{''}, g_{2,s}^{''}, ..., g_{N,s}^{''})$ such that $\sum_n p_{n,s} g_{n,s}^{''} \leq m_s + \sum_n p_{n,s} g_{n,s}$ and $L_s^{''} \succsim_P L_s$.

[72] If states were not mutually exclusive (with exactly one state occuring), it would be hard to believe that properties of other states would not cause preference reversals. For example, one might prefer a 100% chance of a chocolate to an 100% chance of a piece of marshmallow. But if we had a 100% chance of marshmallow with an additional 50% chance of chocolate, this may now be preferred to a 100% chance of a chocolate with an additional 50% chance of additional chocolate.

[73] This assumption of linear pricing is for ease of notational simplicity and could be instead considered as a vector function.

[74] In this case, because the monetary good is allowed to enter directly into the bundle, market prices may influence preferences over bundles.

In words, Monetary Equivalence (i) states that in a case with no uncertainty, the agent is at least as happy off with converting any particular bundle into the money it would cost to purchase that bundle. Since this is true for all degenerate lotteries, including optimal bundles, it also implies that there are no transaction costs to converting money into goods.

Monetary Equivalence (ii) states that in a case with no uncertainty, the agent has no particular preference for holding onto money. In other words, money is only as useful as the things it can buy.[75] It is also worth noting that this does not mean that every dollar must get spent in an optimal bundle. For example, if goods are discrete rather than continuous, it may not be optimal to spend every last dollar. However, what this assumption indicates is that any money left over after purchasing the optimal goods bundle would have no value (as they would be indifferent between that and the same goods bundle with no money).[76]

Independence Property. For any lotteries L and $L^{'}$ in $\mathbb{L}_R$, preferences are independent if $L \succsim_P L^{'}$ implies $\forall \alpha \in (0,1)$ and for all degenerate lotteries $L^{''}$, $(\alpha L,\, (1-\alpha)L^{''}) \succsim_P (\alpha L^{'},\, (1-\alpha)L^{''})$.

With monetary equivalence under uncertainty and the independence property, we can establish the following: Let $G_L = (g_{1,1}, g_{2,1}, ..., g_{N,1}, g_{1,2}, ..., g_{N,2}, ..., g_{1,S}, g_{2,S}, ...g_{N,S}) \in G_1 \times G_2 \times ... \times G_N \times G_1 \times ... \times G_N$ denote the vector of goods for a given lottery L in which all monetary values are 0. Let $G(P,\, I) = \{G_L\, s.t. \sum_s \sum_n p_s g_{n,s} \leq I\}$, with $sup\, G(P,\, I)$ defined using the partial ordering $\succsim_P$ in which all monetary values are set to 0. By a

[75]Although this model is being presented as a static one, the same item at different periods could be thought of as different goods – as long as the uncertainty is resolved in one period with discrete and finite time periods, the same results hold true intertemporally.

[76]For example, let's say I am buying discrete apples and bananas. If apples are \$2 and bananas are \$3 and I have \$7 to spend, I may indeed prefer 2 bananas even though I have \$1 left over. But according to Monetary Equivalence (ii), I am indifferent between \$0 and 2 bananas and \$1 and 2 bananas.

similar notation, let $M_L = (m_1, m_2, ..., m_S) \in \mathbb{R}_S$ denote the vector of monetary values for a lottery L in which all non-monetary goods are 0. And $M(P,\, I) = \{M_L\, s.t.\, \sum_s m_s \leq I\}$, with $sup\, M(P,\, I)$ defined using the partial ordering $\succsim_P$ in which all non-monetary goods are set to 0.

Theorem (Monetary Equivalence Over Uncertainty): Under the assumptions of Monetary Equivalence Under Certainty and Independence, if a lottery of goods is optimal, then the monetary lottery (with equivalent value in each state) will also be optimal. Mathematically, if $G^* = (g_{1,1}, g_{2,1}, ..., g_{N,1}, g_{1,2}, ..., g_{N,2}, ..., g_{1,S}, g_{2,S}, ...g_{N,S}) \in sup\, G(P,\, I)$, then $M^* = (p_{11}g_{11} + p_{21}g_{21} + ... + p_{N1}g_{N1}, p_{12}g_{12} + p_{22}g_{22} + ... + p_{N2}g_{N2}, ..., p_{1S}g_{1S} + p_{2S}g_{2S} + ... + p_{NS}g_{NS}) \in sup\, M(P,\, I)$.

Proof: For proof by contradiction, assume that the condition is true, that $G^* \in sup\, G(P)$ but that, as defined above, $M^* \notin sup\, M(P,\, I)$. Note that G^* can be rewritten as the combination of degenerate lotteries $G^* = \gamma_1 L_1^* + \gamma_2 L_2^* + ... + (1 - \sum \gamma_s)\, L_S^*$ where $L_s^* = (1, 0, g_{1,s}^*, g_{2,s}^*, ..., g_{N,s}^*)$. Individually, each of these degenerate lotteries is weakly dominated by the degenerate lottery $L_s^{'} = (1, p_{1,s}g_{1,s}^* + p_{2,s}g_{2,s}^* + ... + p_{N,s}g_{N,s}^*, 0, ..., 0)$ via Monetary Equivalence under Certainty. By multiple applications of the independence assumption, this means that $(\gamma_1 L_1^*,\, \gamma_2 L_2^*, ..., \gamma_S L_S^*) \precsim_P (\gamma_1 L_1^{'},\, \gamma_2 L_2^{'}, ..., \gamma_S L_S^{'})$ but note that this compound lottery corresponds precisely to M^*.

However as $M^* \notin sup\, M(P,\, I)$ but $M^* \in M(P,\, I)$, that implies there is some $M^{**} \in M(P,\, I)$ with $M^{**} \succ_P M^*$. We can rewrite this lottery as a combination of degenerate lotteries $(\gamma_1, m_1^{**}, 0, ..., 0; \gamma_2, m_2^{**}, 0, .., 0; ...; \gamma_S, m_S^{**}, 0, ..., 0) = (\gamma_1 L_1^{**},\, \gamma_2 L_2^{**},\, ...,\, \gamma_S L_S^{**})$. However, for each of these degenerate lotteries, L_s^{**} the Monetary Equivalence Under Certainty property (ii) states that there exists a degenerate lottery $L_s^{''} = (1, 0, g_{1,s}^{**}, g_{2,s}^{**}, ..., g_{N,s}^{**})$ such that $L_s^{''} \succsim_P L_s^{**}$. Repeated application of the Independence property gives us $G^{''} \equiv (\gamma_1 L_1^{''},\, \gamma_2 L_2^{''},\, ...,\, \gamma_S L_S^{''}) \succsim_P (\gamma_1 L_1^{''},\, \gamma_2 L_2^{''},\, ...,\, \gamma_S L_S^{''})$. Thus $G^{''} \succsim_P M^{**} \succ_P M^* \succsim_P G^*$. This is a contradiction, however, as G^* was the supremum of $G(P,\, I)$ and now there is a new lottery $G^{''}$in $G(P,\, I)$ which strictly dominates it.

2.2.1. Discussion

The assumptions that drive the theory in this case warrant additional discussion. First, if the preference relation is a weak order, that implies that the agent has both transitive and complete preferences. Transitivity of preferences over risk has been discussed as early as Tversky (1969) but more recent empirical evidence suggests that preferences can largely be summarized as transitive (c.f. Birnbaum and Gutierrez (2007); Birnbaum and Schmidt (2010); Regenwetter et al. (2011)).[77] Completeness of preferences is harder to test, as indecision between two lotteries might be interpreted as indifference. This is especially difficult given the great number of goods available on Amazon.com.

Regarding Monetary Equivalence under Certainty, part (i) states that the agent would be at least as well off with an equivalent amount of money as a goods bundle would cost. However, if agents are somewhat unaware of the goods available or the prices of the goods, this may not be the case. The agent might not remember that a good is available to purchase, or the agent might have incorrect beliefs about what the prices, influencing money is capable of purchasing. Indeed, this may be a big component for why individuals seem to treat money and goods differently. To minimize these concerns, Amazon.com was employed for the study, which as a real market should mimic the trade off between money and goods.[78] Lastly, an additional information treatment was employed to test this theory as is described in Section 3.

Monetary Equivalence under Certainty part (ii) states that money holds no inherent value above and beyond what can be purchased with it. In other words, with a given amount of money, the agent can always find a bundle that makes them at least as happy. Yet this assumption makes no mention of the psychic costs that may be associated with finding the bundle in question. In addition, this assumption may be true in our static model and the (static) experiment, but intertemporally agents may want to hold on to some of their money

[77] However, this is an ongoing field of research. It is also possible that the research may not apply to the lotteries employed, being arguably more intricate than previous lotteries studied. Yet intransitive preferences would also make choosing a bundle more difficult in this setting. In previous studies, the outcomes in particular states were largely fixed, whereas in this case the outcomes are subject-determined.

[78] In addition, subjects were allowed to sign in to Amazon if they preferred, to view existing wishlists.

as future prices are not perfectly known.[79]

The Independence property is similar to the Independence property assumed for von Neumann-Morgenstern utility functions. In that set up, lotteries are probability distributions over fixed outcomes. If all goods are discrete, then as the possible lotteries are bounded by the endowment income, then the lottery structure in section 2 could be rewritten under that framework,[80] and the Independence property would be identical.

However the Independence property has been criticized as potentially too strong an assumption. In particular, the famous Allais 'paradox' in which the chance of another lottery may cause preference reversals. Yet the relative importance and frequency of these non-independent lotteries for decision making is an ongoing debate (c.f.Rubinstein (1988); Allais and Hagen (2013)).

2.3. Experiment Design Overview

In order to test this theory, I conducted an incentivized experiment with 124 undergraduate students at the Wharton Behavioral Lab in March 2016. During this experiment, the subjects selected goods on Amazon.com over uncertain states. Subjects sit at the computer and are informed about the upcoming uncertainty. Depending on their treatment, they select either Amazon.com credit (monetary allocations) or Amazon.com goods over several possible states. A wide variety of goods are available on Amazon.com,[81] making this an ideal environment for measuring risk preferences over goods and money.

The static decision is a 2x2x2 design, with agents allocating either {credit or goods} worth a total of {$20 or $100} and is {required or not required} to spend an extra 5 minutes

[79] If future prices were perfectly known, then a good in different time periods could enter the model as different goods. However in addition to price uncertainty, there may be quantity uncertainty, e.g. car stolen, that might make Monetary Equivalence under Certainty (ii) unlikely to be true over time.

[80] If all goods are discrete and bounded by the endowment, then with a finite number of goods and states, there would only be a finite number of possible bundles. As a result, one could rewrite every possible goods bundle in every different state as a different (fixed) outcome. Doing so would make the intuition behind Monetary Equivalence under Certainty hard to understand, and the assumption of discrete goods is unnecessary to the proof I outline.

[81] As of 2015, it is estimated that there are between 300 and 400 million unique items sold on Amazon.com.

browsing Amazon.com. Subjects perform this procedure twice (two rounds).[82] When given a total of \$20, only 2 (equal probability) states can occur, but to keep the average payout the same, 10 (equal probability) states can occur when given a total of \$100.

For example, they might be given \$20 of Amazon.com credit to allocate over two states, A or B, each of which occurs with 50% probability. In this case, a typical "risk averse" decision would be allocate \$10 of credit for State A and \$10 of credit for State B, thus ensuring that regardless of which state occurs, \$10 of Amazon.com credit will be selected. They are then required to spend any credit rewarded.

Alternatively, the subject might be given \$20 of Amazon.com credit, but rather than asked to allocate the credit, the subject selects Amazon.com goods whose prices add up to at most \$20. In other words, the subject determines what to "spend" the credit on goods before the uncertainty is resolved.

Important to this interpretation is the intertemporal fungibility of the Amazon.com credit. Amazon.com goods are purchased at one point in time. Thus, it is important to limit Amazon.com credit to a similar (static) time period. To test the theory outlined above, subjects were informed and quizzed that no matter what credit amount is selected, a single item would be selected at the end of the session whose price is less than or equal to the amount of credit. [83]

To remove concerns about "shrouded attributes" (c.f. Gabaix et al. (2006); Chetty et al. (2009); Brown et al. (2010)), only the list price of the good is considered. Subjects are informed and quizzed in both cases that only the list price will count toward the total, not shipping. In addition, for any URL entered, the browser instantaneously used the Amazon Affiliate API to calculate the price of the item. At the same time, a "total counter" at the bottom of the page informed subjects about the remaining credit available. Combined, these measures aim to limit any "price uncertainty" to prices of unsearched items, rather

[82] Every agent receives both \$20 or \$100 treatments, but it is randomized which occurs first.

[83] 2 subjects selected physical Amazon.com gift cards using either used their Amazon.com credit or Amazon.com items. This was not explicitly discouraged, as an individual willing to do this has made both the Amazon.com goods and Amazon.com credit fungible. However, dropping these individuals makes no difference to the qualitative results or significance.

than the prices of items already selected or searched. For example, I may not know the precise prices of oven mitts, but once I find a particular oven mitt on Amazon.com, all price uncertainty of that specific oven mitt should be gone. Without taxing or shipping concerns, there are no further mental calculations required.[84]

A sample of subjects were randomly selected to spend an extra 5 minutes on Amazon.com to help understand the potential for price or product uncertainty driving potential results. This will be discussed in more detail in Section 4.

Prior to being allowed to start each period, the subjects had to correctly answer questions about the upcoming period, as seen in Appendix Figures 1 and 2. These procedures were implemented to ensure subjects fully understood the incentives they faced. To remove any subject overlap, the computer cookies and browsing history were also cleared in between sessions.

2.4. Experiment Results

The first question is whether the primary treatment of selecting Amazon.com goods (rather than credit) impacted the distribution of good value. Recall that under the assumptions of section 2, there should be no difference between the good distributions and the monetary distributions. For example, if the agent preferred a 10% chance of a $100 item to a sure thing of a $10 item, then when selecting monetary distributions, they should have also preferred a 10% chance of $100 credit to a sure thing of $10 credit. As the credit needed to be spent immediately after awarded, there are no intertemporal savings, so any difference in distribution over the uncertain states would indicate one of the assumptions was not satisfied.

Result 2.4.1 *Contrary to the equivalence theory presented, subjects exhibited greater risk aversion when selecting credit amounts than they did when selecting goods. When selecting goods, subjects were also significantly less likely to select goods of the same value.*

[84] To further simplify things, agents are informed that only the "default" seller price matters. This is primarily Amazon.com itself.

There are several ways to analyze these differences in distribution. I provide results using multiple methods, including regressions of the standard deviation, tests of allocating risk equally across all states, and nonparametric methods to test differences in distribution. All of these methods support the conclusion that subjects selecting credit allocations were more likely to spread out the total amount over multiple states, while subjects selecting goods chose more risky allocations (measured with the price of the items).

Specifically, when investigating the standard deviation of values selected across the possible states, distributing credit meant that agents reduced the standard deviation of the distribution by a third. As seen in Table 3A, OLS estimates suggest the standard deviation of prices were significantly reduced by -2.66 to -2.43 down from a mean of 7.61. However, further analysis of the allocations suggests that not only the standard deviation, but also the mean differs between credit and goods treatments. This may be because goods on Amazon.com are discrete – it is likely difficult or suboptimal to spend precisely \$20 (or \$100).[85] As a result of this discreteness, one might also want to investigate the standard deviation after normalizing values by the total amount allocated. However, the results in Table 3B are nearly identical, with one-third of the standard deviation decreasing when selecting credit.[86]

In the experiment, 16% of subjects removed all risk by allocating a uniform distribution (equal values across all of the uncertain states). As seen in Table 4, this risk-less distribution were over 4 times as likely to occur when the subjects were selecting credit than when they were selecting goods ($p < 0.01$). Note that subjects were instructed and quizzed that they could place the same the same good in multiple slots, thus increasing the chance of having it selected (see Appendix Figure 2). Despite this quizzing, there were only 8 cases where a subject selected a uniform distribution of good prices, indicating a greater tolerance for risk.

[85] Indeed, the average allocation is \$9.32 instead of the full \$10. Selecting credit instead of the goods increases this average by about \$0.50 (more details in Appendix Table 1).

[86] It is worth noting that the \$100 treatment effect changes sign. This is likely because with more money to spend the potential standard deviation of allocations can increase; but once normalized to fractions, this effect goes away.

In addition to these regression results, one can also non-parametrically analyze the distribution of values. However, these tests assume independence among observations, so it is not possible to simply use every individual allocation datapoint. Instead, each allocation is transformed into a single variable that can then be non-parametrically tested across the two primary treatments (goods and credit). The \$20 treatment is a good starting point for this, as most of the information of an allocation can be summarized in a single number, specifically "What is the price of the lower-priced good?" These distributions across individuals are plotted in Figure 2A, and the associated Kolmogorov-Smirnov test rejects equality of the distributions ($p < 0.01$). Figure 2B plots distributions of a similar nature, that is the normalized price of the lower-priced good (in other words, what fraction of the total spent is on the lower-priced good). Kolmogorov-Smirnov suggests borderline significant rejection for equality of the distributions of this transformation ($p < 0.07$).

However, though widely used, the Kolmogorov-Smirnov test uses the largest difference between the distributions. As a result, it tends to underweight differences in the tails of the cumulative distributions – Mason and Schuenemeyer (1983); Kim and Whitt (2015). Given the large share of subjects who place \$10 and \$10 when using credit, the Kolmogorov-Smirnov test may not be the most efficient. Alternatively, we can also use more information from the distributions, such as a Kolmogorov-Smirnov test of the within-allocation standard deviation. In these cases, the Kolmogorov-Smirnov rejects equality of distribution both when using the distributions of standard deviations ($p < 0.01$) or the distributions of normalized standard deviations ($p < 0.01$).

Result 2.4.2 *When forced to spend more time searching Amazon.com, subjects did not significiantly alter the distribution allocations of goods and credit.*

To test the possibility that the difference in risk for the good domain is being driven by product or price uncertainty, some subjects were randomly submitted to an information treatment. In this treatment, subjects were forced to wait an extra 5 minutes before they could submit their allocations. During this time, subjects were only allowed to visit Amazon.com or sit quietly at the desk.[87] The intent was to lower the marginal cost of searching. It appears this treatment was indeed successful in inducing subjects to spend more time in a section – the average treatment effect was to spend an extra 8 minutes (3 minutes beyond the 5 minutes imposed). This extra time spent searching could be the result of product search being unexpectedly fun or that the the 5 minute timer was not visible while browsing Amazon, causing subjects to run over.

As we can see in the OLS regressions in Table 5A, the treatment information had no significant direct impact on allocation distributions (as measured by the standard distribution). If the information treatment's effect on allocation would be through the time spent searching, we can also use the information treatment as an instrumental variable for time spent in a section. This allows a causal impact of time spent searching on the distribution allocations. Table 5B presents results of this instrumental variable regression, but once again, spending more time searching has no significant impact on the standard deviation of the allocation.

2.5. Conclusion

Contrary to the equivalence theory money and goods under uncertainty, subjects exhibited reduced risk taking with selecting credit amounts than they did when selecting goods. When selecting goods, subjects were also significantly less likely to select goods of the same value across the uncertain states. These findings alone might indicate a general uncertainty of Amazon.com goods or prices, but forcing subjects to spend more time investigating Amazon.com does not change these differences.

As a result, one of the remaining assumptions of the equivalence theory must be false to result in this behavior. One possibility is the an endowment effect, which allows for ownership of the item to increase the owner's willingness-to-accept (Knetsch (1989); Kahneman et al. (1991)). In this experiment, if committing to a good allocation triggers a similar endowment effect, then a post-committed good allocation might be worth more than an equivalent price bundle. In order for this to cause greater risk taking with goods, there

[87] This website restriction, as well as a no cellphone rule, was enforced by lab assistants monitoring the study.

must be convexities in the endowment effect; otherwise good bundles would not necessarily be more risky.

Another possibility is a model of thinking aversion presented in Ortoleva (2013). Although the information treatment resulted in subjects spending more time searching, it may be that they still dislike the large choice set. As a result, they may not decide on a particular allocation of goods until they have no other choice. This could result in less risk taking in credit relative to goods.

Lastly, it may be that credit, being easier to compare, may result in more "regret aversion" as described in Loomes and Sugden (1982). If the prices of goods or inherent value of goods makes comparisons more difficult, it may be the case that the agent will experience less regret if the "best" outcome does not happen. As a result, they may be more willing to engage in riskier behavior over goods than the more easily comparable money outcomes.

It may also be the case that by reducing the choice set, the difference between money and goods risk taking could decrease. For example, if subjects were only able to choose between two goods for each uncertain state, one might expect a convergence of money and good risk taking. But in the real world, individuals face many possible uses for their money.

While these are interesting possibilities and warrant further study, this does not change the primary finding of this paper. In other words, whether convex endowment effects or thinking aversion is driving the difference in risk taking, it remains that individuals react differentially to risk over goods and risk over money.

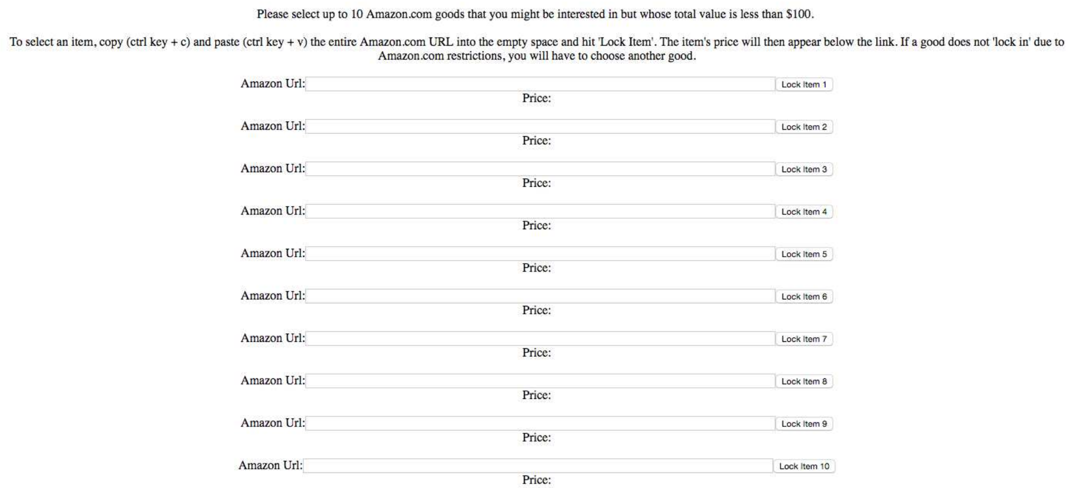
This finding has important implications for public policy. In 2014 U.S. government sponsored lotteries raised $70 billion in revenues, helping fund state governments and programs. This paper suggests that individuals may be more willing to engage in lotteries that have goods, not just money. Indeed, anecdotally U.S. companies often run sweepstakes with prizes (cars, cruises, etc.) rather than a pure lottery.[88] For example, a prize of $1 million with a car worth $50,000 may cause more engagement in risk than a lottery with $1,050,000

[88]Some of this is likely driven by reduced costs of prizes, which may be seen as a marketing cost. However, this reduced cost may also be achieved for a potential state run sweepstakes.

million. Although the total ramifications of state-run lotteries are debatable, this greater willingness in risk could be used to reduce advertisement and overhead budgets without changing revenues.

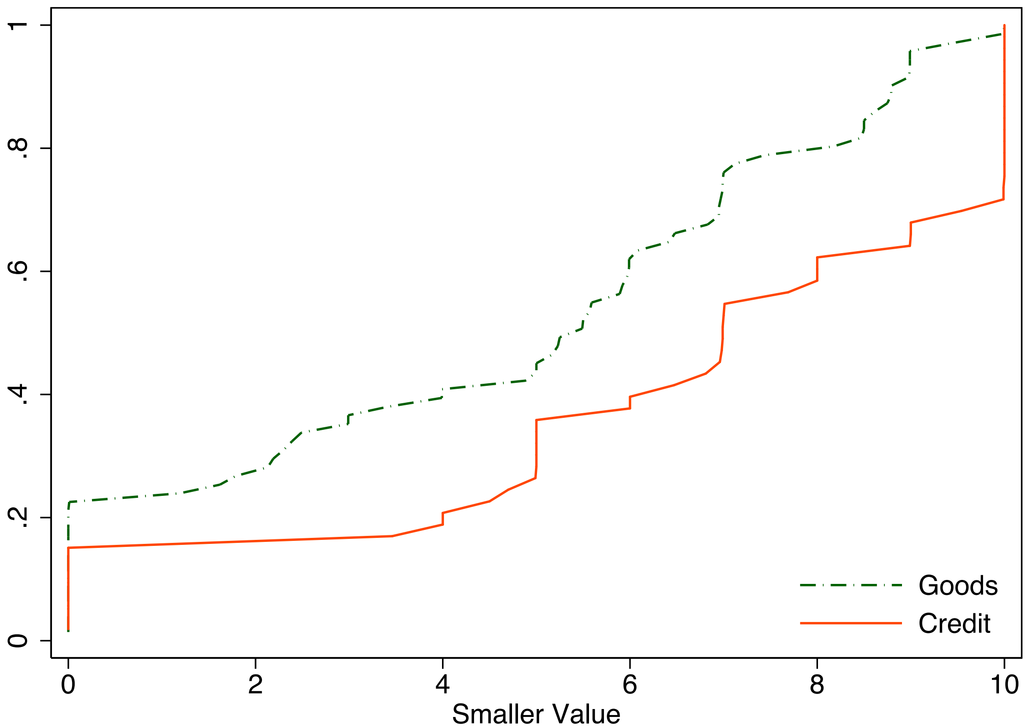
2.6. Figures

Figure 6: Example of Good Selection



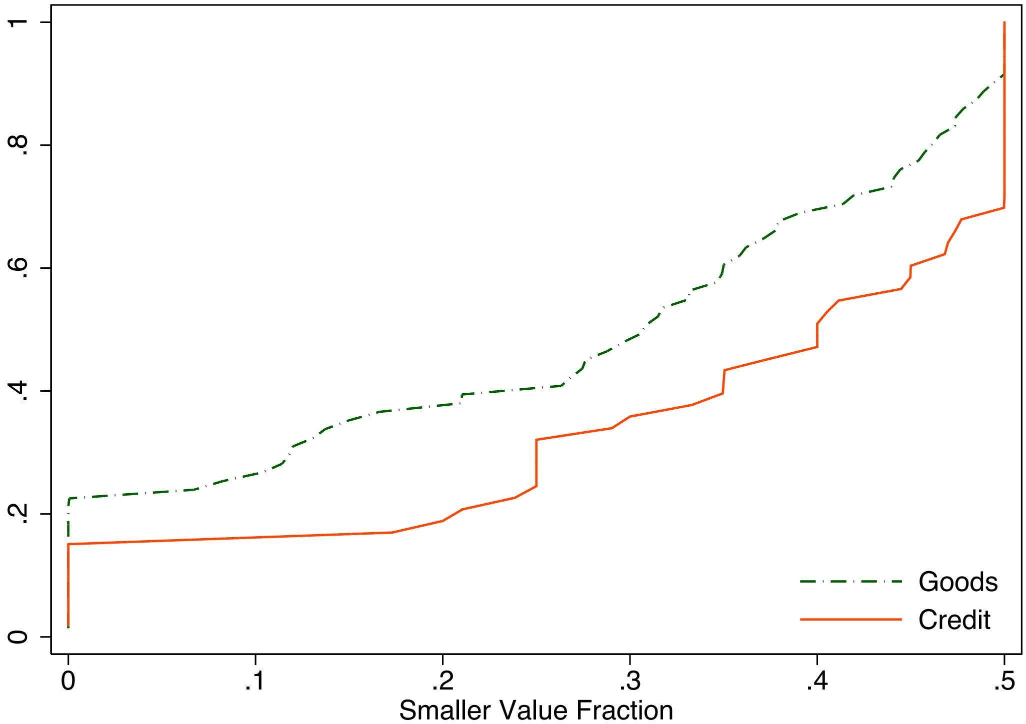
Notes. Figure demonstrates a typical good selection screen faced by subject. Whether subject was asked to select Amazon.com goods (via URLs) or Amazon.com credit amounts was randomized. Whether subject was asked to find up to 10 items that totaled at most $100 or up to 2 items that totaled at most $20 was also randomized. See Experiment Design for more details.

Figure 7: Distribution of the Smaller Value When Total is $20



Notes. Plot shows two cumulative distributions of the value of the smaller good when total is $20. When given $20 to allocate, the agent chooses to allocate across 2 uncertain states, this represents the smaller of these two allocations. Vertical axis represents the frequency of that value occurring across the two different treatments (selecting goods or selecting credit allocations).

Figure 8: Distribution of the Smaller Value (Normalized) When Total is $20



Notes. Plot shows two cumulative distributions of the (normalized) value of the smaller good when total is $20. Value is normalized by dividing by the total value allocated. When given $20 to allocate, the agent chooses to allocate across 2 uncertain states, this represents the smaller of these two allocations. Vertical axis represents the frequency of that (normalized) value occurring across the two different treatments (selecting goods or selecting credit allocations).

2.7. Tables

Table 16: Summary Statistics

	Mean	Standard dev	Min	Max	
Individual Level Variables					
Female	0.71	0.45	0	1	
Age	20.2	1.3	18	24	
SAT Math Score	733	62	540	800	(22 missing)
Computer Skill Test	2	0	2	2	(1 missing)
Number of Previous Lab Studies	26.6	24.8	1	133	
Period Level Variables					
Average Value of Entry	$9.33	1.07	4	10	
Standard Dev of Entry (within)	$7.06	6.92	0	31.6	
$100 Treatment Indicator	0.50	0.50	0	1	
Credit Treatment Indicator	0.46	0.50	0	1	
Time Spent Searching (seconds)	487	373	45	1699	
Number of Individuals	124				
Number of Treatment Periods	248				

Notes. Computer Skill Test was a demographic variable collected by the Wharton Behavioral Lab prior to the experiment, however among subjects above it had no variation. SAT Math score is missing for individuals who either took the ACT or otherwise did not wish to share that information with researchers.

Table 17: Randomization Check

Dependent Variable	Credit Treatment		Period # for $100 Treatment	
Female	−0.11	−0.01	−0.02	−0.01
	(0.07)	(0.11)	(0.10)	(0.11)
SAT Math Score		0.01		0.01
('00s of points)		(0.08)		(0.09)
Previous WBL Studies		−0.001		−0.001
		(0.002)		(0.002)
F-test	2.59	0.24	0.05	0.24
p value	0.11	0.87	0.83	0.87
Dependent Variable Mean	0.46	0.48	1.54	1.54
Number of Observations	248	204	248	204
Number of Individuals	124	102	124	102

Notes. Standard Errors (clustered at individual level) presented in parentheses above. As every subject in experiment 1 receives both the $20 and $100 treatments, the dependent variable for $100 treatment is the period in which they received the treatment in question. If randomization was done properly, the pre-treatment variables should not predict the period they received this treatment. Indeed, the F-stats are all large enough that I fail to reject the hypothesis that all coefficients are zero under $\alpha = 0.05$. Thus, I conclude the randomization was adequately done. SAT Math score is missing for 22 individuals who either took the ACT or otherwise did not wish to share that information with researchers.

Table 18: Credit and \$100: Impact on Standard Deviation of Selection Value

Dependent Variable:	Specification			
Value Standard Deviation	(1)	(2)	(3)	(4)
Subject Selects Credit	−2.66***	−2.61***	−2.56***	−2.43***
(Binary Treatment Var.)	(0.77)	(0.77)	(0.81)	(0.82)
\$100 Total Allocation	4.96***	5.00***	5.00***	4.99***
(Binary Treatment Var.)	(0.64)	(0.63)	(0.64)	(0.64)
First Period		0.59	0.59	0.61
		(0.64)	(0.66)	(0.67)
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	7.61	7.61	7.61	7.61
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.16	0.16	0.19	0.22

Notes. The dependent variable is the standard deviation of value of the entries in a single period. All specifications report results from OLS regressions and also include a constant term. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 19: Credit and \$100: Impact on Standard Deviation of Selection Value (Normalized)

Dependent Variable:	Specification			
Normalized Value Std Dev	(1)	(2)	(3)	(4)
Subject Selects Credit	−0.08***	−0.07***	−0.07***	−0.07***
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
\$100 Total Allocation	−0.17***	−0.17***	−0.17***	−0.17***
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
First Period		0.01	0.01	0.01
		(0.02)	(0.02)	(0.02)
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	0.20	0.20	0.20	0.20
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.20	0.20	0.25	0.27

Notes. The dependent variable is the standard deviation of normalized value of the entries in a single period. Values were normalized by dividing by the total value allocated. All specifications report results from OLS regressions and also include a constant term. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 20: Credit and \$100: Impact on Equality of Selection Values

Dependent Variable:	Specification			
All Entries Same Value	(1)	(2)	(3)	(4)
Subject Selects Credit	0.22***	0.21***	0.23***	0.21***
(Binary Treatment Var.)	(0.05)	(0.05)	(0.04)	(0.04)
\$100 Total Allocation	−0.07	−0.07*	−0.08*	−0.07*
(Binary Treatment Var.)	(0.04)	(0.04)	(0.04)	(0.04)
First Period		−0.13***	−0.13***	−0.13***
		(0.04)	(0.04)	(0.04)
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	0.18	0.18	0.18	0.18
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.10	0.13	0.21	0.22

Notes. The dependent variable is the whether all entries have equal value in a single period. Values were normalized by dividing by the total value allocated. All specifications report results from OLS regressions and also include a constant term. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 21: OLS Impact of Search on Standard Deviation of Selection Value

Dependent Variable:	Specification			
Normalized Value Std Dev	(1)	(2)	(3)	(4)
Subject Selects Credit	−0.08***	−0.08***	−0.07***	−0.07***
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
$100 Total Allocation	−0.17***	−0.17***	−0.17***	−0.17***
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
Information Treatment	0.01	0.01	0.03	0.04
	(0.02)	(0.02)	(0.03)	(0.03)
Round Fixed Effects		X	X	X
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	0.20	0.20	0.20	0.20
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.21	0.21	0.25	0.27

Notes. The dependent variable is the standard deviation of normalized value of the entries in a single period. Values were normalized by dividing by the total value allocated. All specifications report results from OLS regressions and also include a constant term. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Table 22: IV Impact of Search on Standard Deviation of Selection Value

Dependent Variable:	Specification			
Normalized Value Std Dev	(1)	(2)	(3)	(4)
Subject Selects Credit	-0.07^{***}	-0.07^{***}	-0.06^{***}	-0.07^{***}
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
\$100 Total Allocation	-0.17^{***}	-0.17^{***}	-0.17^{***}	-0.17^{***}
(Binary Treatment Var.)	(0.02)	(0.02)	(0.02)	(0.02)
Time Searching (Minutes)	0.001	0.001	0.003	0.004
	(0.003)	(0.003)	(0.003)	(0.003)
First Stage F Stat (IV)	176.9	168.8	159.2	144.2
Round Fixed Effects		X	X	X
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	0.20	0.20	0.20	0.20
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.21	0.21	0.25	0.27

Notes. The dependent variable is the standard deviation of normalized value of the entries in a single period. Values were normalized by dividing by the total value allocated. All specifications report results from GMM Instrumental variable regressions and also include a constant term. Time spent searching was instrumented by the information treatment, with F values from the first stage reported. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

CHAPTER 3 : Time Lotteries

3.1. Introduction

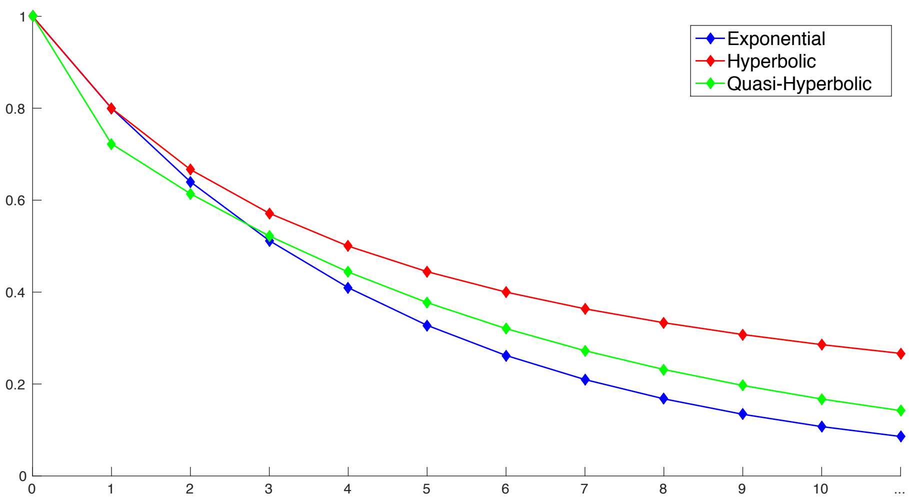
Suppose you are offered a choice between *(i)* receiving $\$x$ in period t for sure, or *(ii)* receiving $\$x$ in a random period $\tilde{t}$ with mean t. For example, you can receive \$100 in 10 weeks for sure, or \$100 in either 5 or 15 weeks with equal probability. Both lotteries pay the same amount and have the same expected delivery date. However, the delivery date is known in the first lottery and is uncertain in the second one. Suppose also that if you choose option *(ii)*, the exact payment date will be revealed immediately, eliminating any planning issues. Which would you choose?

In this paper we study preferences over time lotteries: lotteries that pay the same prize at uncertain future dates, with the uncertainty fully resolved immediately after the choice is made.[89] Uncertainty about timing has a bearing on many real life choices. For example, when one decides whether to invest in a project that is certain to start paying dividends in 5 years rather than in one which starts payments within an *average* of 5 years. Another example is whether to pay more to guarantee that the delivery date of an online purchase (like a book from Amazon) is t days from now, rather than t days on average.

The starting point of our analysis is the observation that the ubiquitous model of time preferences – Discounted Expected Utility with convex discounting – imposes a specific direction of these preferences: independently of the values of x and t, when offered a choice between options *(i)* and *(ii)* above, subjects should *always* pick the option with an uncertain payment date *(ii)*; that is, they must be *risk seeking over time lotteries* (RSTL). To see this, recall that the standard way to evaluate a consumption path that pays c in each period and $c+x$ in some period t is given by $\sum_{\tau \neq t} D(\tau)\, u(c) + D(t)\, u(c+x)$, where u is a time-independent utility function over outcomes and D is a decreasing discount function. If the

[89]The notion of time lotteries is introduced in Chesson and Viscusi (2003) (termed "lottery-timing risk aversion"), and also analyzed in Onay and Öncüler (2007); we discuss the relation with these papers below.

Figure 9: The exponential, hyperbolic, and quasi-hyperbolic discount functions



Notes. The exponential, hyperbolic, and quasi-hyperbolic discount functions are, respectively, $D(t) = \delta^t$, $D(t) = \frac{1}{1+\gamma t}$, and $D(t) = \beta\delta^t$ for $t \geq 1$ (with $D(0) = 1$), where $\beta, \delta \in (0,1)$ and $\gamma > 0$. They are all convex.

consumption path is random, overall utility is obtained by taking expectations, leading to the Discounted Expected Utility (DEU) model. As long as the discount function D is convex, Jensen's inequality implies that option *(ii)* is preferred. Note that this is independent of the curvature of the utility function u, as the payment is the same in both options. Since virtually *all* discount functions used in economics are convex (including exponential, hyperbolic, and quasi-hyperbolic; see Figure 9) and since non-convex discounting is possible at only finitely many points (as we discuss in Section 3.2), this is a fundamental feature of the standard model.

We contend that this prediction of DEU is too strong: in an incentivized experiment, we find that most subjects are not RSTL. Instead, subjects normally pick lotteries with known payment dates, exhibiting risk aversion over time lotteries (RATL). Our theoretical contribution is to show that two well-known generalizations of the standard model — the separation of risk aversion and time preference in Epstein and Zin (1989) (henceforth EZ) and a discounted non-Expected Utility model with probability weighting — can both account for this behavior. We then use our experimental findings to demonstrate that EZ better fits the data. Our results provide a new justification for the separation of attitudes toward risk and intertemporal substitution in EZ and also suggest new theoretical restrictions on its parameters.

According to EZ, time lotteries are evaluated by first computing the discounted utility of each consumption path in its support. Then, unlike in DEU, these discounted utilities are aggregated *non-linearly*; the certainty equivalent of the lottery over them is computed in a way that depends on the curvature of a function v, which captures the individual's risk aversion and is different from the function u used in evaluating deterministic consumption paths. If the individual is risk averse enough (v is sufficiently concave), then she will prefer a payment in a known date to a payment with an uncertain but mean preserving date. Thus, EZ predicts a correlation between risk aversion over time lotteries and standard atemporal risk aversion, since both are affected by the same parameter.

Another way to accommodate some risk aversion over time lotteries, also suggested in Onay and Öncüler (2007), is to relax Expected Utility by aggregating the discounted utilities using expectation with respect to distorted probability weights. This Discounted Probability Weighting Utility model (DPWU) predicts a preference for known payment dates only among subjects who violate Expected Utility by sufficiently underweighting the probabilities of good outcomes.

We conducted an experiment to test if subjects are risk seeking over time lotteries, and, if not, to separate between the two explanations above. We first asked subjects to choose between pairs of time lotteries, where the distribution of dates in one option was a mean preserving spread of that in the other. We then measured standard time preferences. Lastly, we asked questions on regular risk preferences to measure subject's atemporal risk aversion, as well as violations of Expected Utility and probability weighting. The following are our key findings:

1. Only a small number of subjects (less than 7%) are RSTL and the majority can be classified as RATL. When both options have random payment dates, most subjects still prefer the less risky ones (in the sense of mean preserving spreads), which suggests that they are not simply attracted to certainty.

2. A large majority of subjects (82%) exhibit convex discounting. The above result (1) remains unchanged if we restrict the analysis to these subjects.

3. Consistently with EZ, preferences for known payment dates are strongly related to atemporal risk aversion.

4. In contrast to DPWU, preferences for known payment dates are unrelated to violations of Expected Utility. Subjects who do not underweight probabilities still prefer known payment dates. The same is true if we focus only on those who abide by Expected Utility theory in atemporal risky choices (i.e., we eliminate those who both underweight and overweight probabilities). In addition, a regression analysis shows that the preference for known payment dates is unrelated to the degree of probability weighting.

Overall, our experimental results show the prevalence of RATL (contradicting DEU), and support the generalization based on EZ, but not the one based on probability weighting. This is of particular relevance because EZ is a widely-used model, especially in macroeconomics and finance. The common justification for adopting EZ is that it allows for two different parameters to govern attitudes toward risk and intertemporal substitution – an additional degree of freedom that has proved particularly effective in matching empirical data.[90] Behaviorally, it is sometimes justified as it allows for a preference for early rather than late resolution of uncertainty. Based on our results presented above, we suggest an additional reason to adopt this model: it permits a wider range of preferences over time lotteries, allowing decision makers to prefer payments with a known date.

[90] See, for example, Bansal and Yaron (2004) and Chen et al. (2013).

The impossibility of DEU to accommodate different attitudes towards time lotteries can be understood with an analogy to the classic work of Yaari (1987). Within the (atemporal) Expected Utility framework, diminishing marginal utility of income and risk aversion are bonded together. But, as Yaari argues, these two properties are "horses of different colors" and hence, as a fundamental principle, a theory that keeps them separate is desirable. We show an analogous property in a temporal setting. Convex discounting, which is a property of *pure time preferences*, necessarily implies RSTL. There is no fundamental reason why the two notions should be related and, in fact, we find their equivalence even more troubling than the equivalence pointed out by Yaari. This is because while diminishing marginal utility of income and risk aversion relate to two different phenomena, they are both reasonable properties of preferences. In our case, while decreasing willingness to wait (convex discounting) is a plausible behavioral property, supported by our experimental data, it seems that most people are not RSTL.

This paper is related to the literature on the interaction between time and risk.[91] Chesson and Viscusi (2003) introduce the idea of time lotteries, analyze the case of standard DEU with exponential discounting, and argue that uncertainty aversion over outcome timing should be correlated with uncertainty aversion over probabilities of outcomes. They conducted a hypothetical survey on business owners, and found that 30 percent of the subjects dislike uncertainty in the timing of an outcome. Onay and Öncüler (2007) generalize their theoretical result, pointing out that (what we call) RSTL holds in DEU for any convex discounting (their analysis also focuses on the roles of gains and losses). They conducted an un-incentivized survey, with large hypothetical payments, and also found that subjects dislike uncertainty in timing. They link this to violations of Expected Utility due to probability distortions. Unlike these two papers, we show that EZ can also theoretically account for a preference for sure dates. Moreover, we conducted an actual (incentivized) experiment in which we confirmed the prevalence of preferences for known payment dates. When we

[91] See Epper and Fehr-Duda (2015) and references therein.

compared possible explanations, we found support for the one based on EZ. Eliaz and Ortoleva (forthcoming) studied the case in which the payment date is ambiguous, as opposed to risky. They found that the majority of subjects remain averse to this ambiguity, but the proportion is much smaller than that of aversion to ambiguous payments.

Other papers in the literature on the relation between time and risk mostly focus on different issues and do not analyze attitude towards time lotteries. For example, Halevy (2008) suggests a link between present bias (Strotz, 1955) and the certainty effect (Kahneman and Tversky, 1979) based on the idea that the present is known while any future plan is inherently risky. He shows that the two ideas can be bridged, if the individual evaluates stochastic streams of payments using specific non-Expected Utility functionals. Andreoni and Sprenger (2012) challenge a property of DEU, according to which intertemporal allocations should depend only on relative intertemporal risk. They conducted an experiment that focused on common-ratio type questions, applied to intertemporal risk. Their result indicates that subjects display a preference for certainty when it is available, but behave closely in line with the predictions DEU when only uncertain (future) prospects are considered. Andersen et al. (2008), Epper et al. (2011), and Dean and Ortoleva (2015) experimentally studied the relation between time and risk preferences. They found that time preferences are correlated with the curvature of the utility function, as predicted by the standard model; their results on the relation with violations of Expected Utility are mixed.

The remainder of the paper is organized as follows: Section 3.2 formally shows that DEU implies risk seeking towards time lotteries. Section 3.3 introduces the two possible generalizations that can allow for a wider range of behavior — EZ and discounted probability weighting — and discuss their empirical predictions. Section 3.4 presents the experimental design and its results. The appendix includes proofs of the formal results and the experimental details.

3.2. DEU and Attitudes towards Time Lotteries

We assume time to be discrete and identify the set of possible times with the set of natural numbers. A consumption path is a sequence $\mathbf{c} = (c_1, c_2, ..., c_t, ..)$ that specifies consumption in each period $c_t \in \mathbb{R}_+$.

Our analysis focuses on a particular type of lotteries over consumption paths, which we call *time lotteries*. Fix a base level of per-period consumption c. For each monetary prize $x \in \mathbb{R}_+$ and time $t \in \mathbb{N}$, let (x, t) denote the consumption path that pays $c + x$ in period t and c in any other period. A time lottery $p_x = \langle p_x(t), (x, t)\rangle_{t\in\mathbb{N}}$ is a finite probability measure over $\{x\} \times \mathbb{N}$, that is, a lottery that yields an additional payment of x in period t with probability $p_x(t)$.[92] The degenerate lottery that yields (x, t) for sure is denoted by $\delta_{(x,t)}$. Let $\mathcal{P}_x$ denote the set of all time lotteries with payment x. Preferences are defined over $\mathcal{P} := \underset{x\in X}{\cup} \mathcal{P}_x$ and are represented by some function $V : \mathcal{P} \to \mathbb{R}$.

We say that preferences are *risk averse towards time lotteries (RATL)* if, for every $x \in \mathbb{R}_+$ and every $p_x \in \mathcal{P}_x$ with $\sum_\tau p_x(\tau)\,\tau = t$,

$$V\left(\delta_{(x,t)}\right) \geq V(p_x). \tag{3.1}$$

They are *risk seeking towards time lotteries (RSTL)* if the sign in (3.1) is reversed.

According to the Discounted Expected Utility model (DEU), the value of each time lottery p_x is given by

$$V_{DEU}(p_x) = \sum\nolimits_t p_x(t)\left[\sum\nolimits_{\tau\neq t} D(\tau)\,u(c) + D(t)\,u(c+x)\right],$$

[92] To avoid confusion, let us emphasize that time lotteries are fundamentally different from lotteries that pay an uncertain amounts in different periods. For example, the time lottery that pays \$1 in either periods 1 and 3 with equal probabilities is very different from the lotteries that pay \$0 or \$1 with 50 percent chance in periods 1 and 3. With the former, the decision maker always gets exactly \$1, although she does not know when. With the latter, she may get a total of \$0, \$1, or \$2.

where $u : X \to \mathbb{R}_+$ is a continuous and strictly increasing utility function, and $D : \mathbb{N} \to (0, 1]$ is a strictly decreasing discount function.

We say that a discount function is *discretely convex* if D is a convex function when defined, that is, if for all $t_1, t_2 \in \mathbb{N}$ and $\alpha \in (0, 1)$,

$$\alpha D(t_1) + (1-\alpha) D(t_2) \geq D(\alpha t_1 + (1-\alpha) t_2)$$

whenever $\alpha t_1 + (1-\alpha) t_2 \in \mathbb{N}$.

The following proposition establishes the relationship between attitudes towards time lotteries and the convexity of the discount function. (The first part of our result, confined to two prizes coded as "gains," is stated as Hypothesis 1 in Onay and Öncüler 2007.)

Proposition 3.2.1 *Under DEU, preferences are RSTL if and only if D is discretely convex. Moreover, they cannot be RATL.*

Proof First, we show that preferences are RSTL (RATL) if and only if D is discretely convex (concave). (A discount function is discretely concave if $-D$ is discretely convex.) The value of $\delta_{(x,\bar{t})}$ is

$$V_{DEU}\left(\delta_{(x,\bar{t})}\right) = \sum\nolimits_{\tau \neq \bar{t}} D(\tau) u(c) + D(\bar{t}) u(c+x),$$

whereas the value of the time lottery $p = \langle p_x(t), t \rangle_{t \in \mathbb{N}}$ with $\sum_t p_x(t) t = \bar{t}$ is

$$V_{DEU}(p) = \sum\nolimits_t p_x(t) \left[\sum\nolimits_{\tau \neq t} D(\tau) u(c) + D(t) u(c+x)\right]$$

With algebraic manipulations give:

$$V_{DEU}(p) \geq V_{DEU}\left(\delta_{(x,\bar{t})}\right) \Leftrightarrow \left[\sum p_x(t) D(t) - D(\bar{t})\right] [u(c+x) - u(c)] \geq 0,$$

which, because u is strictly increasing, holds if and only if D is convex.

Next, we show that D cannot be discretely concave. Suppose D is discretely concave, so that

$$D(t) \leq D(1) + (t-1)\left[D\left(2\right) - D\left(1\right)\right].$$

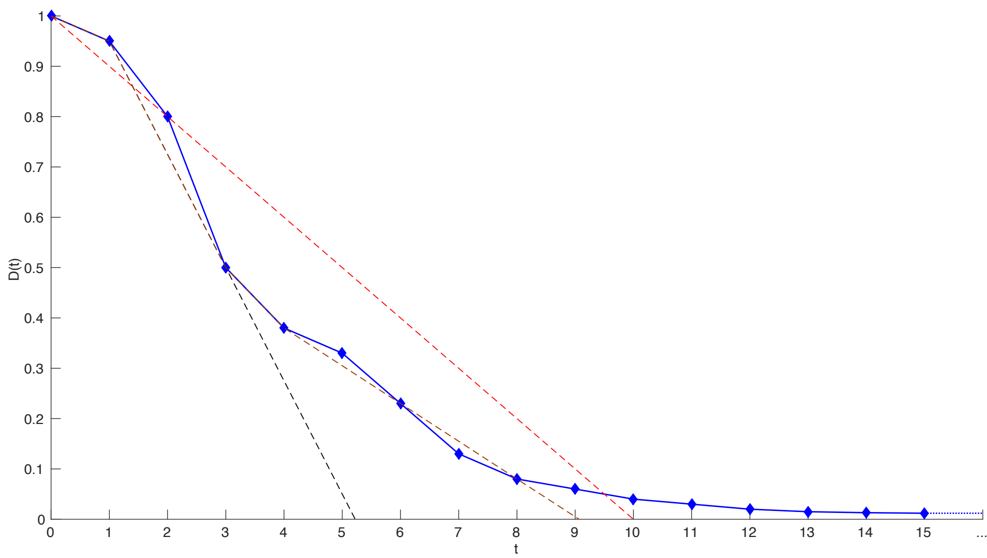
Taking $t \geq \frac{2D(1)-D(2)}{D(1)-D(2)}$ and using the fact that D is strictly decreasing, we obtain $D(t) < 0$, which contradicts the fact that the discount function is positive.

The first part of Proposition 3.2.1 states that if the discount function is convex, a DEU decision maker *must* be RSTL. Importantly, this result does not rely on the curvature of the utility function u since all options involve the same payments, although in different periods. But is it plausible to assume that the discount function is convex? As we pointed out in the introduction, virtually all discount functions used in economics are convex (see Figure 9). Indeed, convexity of the discount function has a natural behavioral interpretation: it means that time delays are less costly the further away in time they occur. Therefore, with standard discount functions, DEU leaves no degrees of freedom on the risk attitude towards time lotteries.

The second part of Proposition 3.2.1 states that discount functions cannot be (globally) concave, implying that we cannot have RATL with DEU. Figure 10 illustrates the intuition behind this result. The discount function must be decreasing. Concavity requires it to decrease at an increasing rate, meaning that the discount function at all periods greater than $t+1$ must lie below the secant passing through $D(t-1)$ and $D(t+1)$. If the discount function were concave, it would eventually cross the horizontal axis, contradicting the fact that discount functions are positive.

In light of Proposition 3.2.1, we ask whether DEU can satisfy weaker, local versions of RATL. Since time lotteries have two dimensions (time and prizes), we consider two different notions of local RATL. We say that preferences are *locally risk averse towards time lotteries at prize*

Figure 10: Discount functions are locally convex in all but a finite number of periods



Notes. Dotted lines are the secants going through the points adjacent to the dates where D is locally concave. To be globally concave, D would have to be below these secants, crossing the horizontal axis.

x if

$$V\left(\delta_{(x,t)}\right) \geq V(p)$$

for every $p \in \mathcal{P}_x$ with $\sum_\tau p_x(\tau)\,\tau = t$. They are *locally risk averse towards time lotteries at time* t if the sure payment at t is preferred to a random payment occurring at either $t-1$ or at $t+1$ with equal probabilities, that is,

$$V\left(\delta_{(x,t)}\right) \geq V\left(\langle 0.5,(x,t-1);0.5,(x,t+1)\rangle\right)$$

for every $x \in \mathbb{R}_+$. As before, we say that preferences are locally risk seeking at either x or t if the reverse inequalities hold.[93]

Proposition 3.2.2 below shows that even these weaker versions of RATL are inconsistent with DEU. Thus, even if we were willing to abandon convexity, it would be of limited

[93] While neither notion of local RATL is contained in the other, being locally RSTL in any of them prevents the decision maker from being locally RATL in the other.

help. Since the previous discussion did not rely on varying the prize x, it follows that DEU also cannot accommodate local RATL at any x. Moreover, as Figure 10 illustrates, local concavity can hold only at a small number of periods. Therefore, preferences are generically locally RSTL in time, in the sense that they are locally RATL in only finite number of dates.

Proposition 3.2.2 *Under DEU, there is no x at which preferences are locally RATL. Moreover, the set of periods in which preferences are locally RATL at t is finite.*

Proof See Appendix.

In the following section, we will see two theories, each of them allowing for one of these notions of local RATL to hold.[94]

3.3. Beyond DEU: Epstein Zin Preferences and Probability Weighting

We now present two generalizations of DEU that allow more flexible attitudes towards time lotteries. The first one disentangles attitudes towards risk from intertemporal substitution, leading to the model of Epstein and Zin (1989). The second replaces objective probabilities by decision weights. We show that both models can capture preferences for known payment dates and discuss their implications. In the next section, we will evaluate each of them using data from our experiment.

3.3.1. Separating Time and Risk Preferences: the model of Epstein Zin

Epstein and Zin (1989) (EZ) introduced a general class of preferences over stochastic consumption paths, defined recursively over this period's known consumption and the certainty equivalent of next period's utility. In the most popular version of EZ, lotteries over consumption paths are evaluated using the recursive formula:

[94]The analysis thus far does not permit individuals to smooth consumption by redistributing their prize over time. In the Appendix, we allow the decision maker to costlessly borrow and save. We show that $\delta_{(x,t)}$ is preferred over $\langle 0.5, (x, t-1); 0.5, (x, t+1) \rangle$ if and only if the decision maker is sufficiently risk averse. However, quantitatively, choosing the safe lottery requires absurdly high risk aversion or extremely large prizes. For example, suppose the (market) rate of discount is 0.9. Then, an individual with a million dollars in discounted lifetime earnings and a constant coefficient of relative risk aversion of 10 would prefer the safe lottery only if the prize exceeds $\$123,500$.

$$V_t = \left\{ (1-\beta)\, c_t^{1-\rho} + \beta \left[E_t \left(V_{t+1}^{1-\alpha} \right) \right]^{\frac{1-\rho}{1-\alpha}} \right\}^{\frac{1}{1-\rho}}, \tag{3.2}$$

where c_t denotes consumption at time t, $\alpha > 0$ is the coefficient of relative risk aversion, and $\rho > 0$ is the inverse of the elasticity of intertemporal substitution. EZ boils down to DEU whenever $\alpha = \rho$. A well-known advantage of this model is that it separates the roles of risk aversion and the elasticity of intertemporal substitution – which must be the inverse of one another with DEU. This additional degree of freedom has proved to be particularly useful in applied work, and this model is widely used in macroeconomics, asset pricing, and portfolio choice. From a behavioral prospective, this generalization of DEU also allows subjects to express preferences for early or late resolution of uncertainty. One message from our analysis below is that separating risk aversion and the elasticity of intertemporal substitution also allows accommodating some risk aversion over time lotteries. This feature, in turns, suggests new theoretical restrictions on the values of the parameters α and ρ.

Given the simple structure of time lotteries, in which all uncertainty about future consumption is resolved immediately, the value of a time lottery $p \in \mathcal{P}_x$ using equation (3.2) is

$$V_{EZ}(p) = \left\{ (1-\beta)\, c^{1-\rho} + \beta \left[E_p \left(V^{1-\alpha} \right) \right]^{\frac{1-\rho}{1-\alpha}} \right\}^{\frac{1}{1-\rho}}, \tag{3.3}$$

where $E_p(\cdot)$ denotes the expectation with respect to the measure p.[95] If we let $\lambda := \frac{c+x}{c} > 1$ denote the proportional increase in consumption from the prize, the continuation utility V is determined by

$$V\left(\delta_{(x,t)}\right) = [(1-\beta)\, c]^{\frac{1}{1-\rho}} \left[\beta^t \lambda^{1-\rho} + \sum\nolimits_{\tau \neq t} \beta^{\tau} \right]. \tag{3.4}$$

Consider a choice between the safe lottery $\delta_{(x,t)}$ and the risky lottery $\langle 0.5, (x, t-1); 0.5, (x, t+1) \rangle$. The next proposition establishes that EZ can accommodate a preference for the safe lottery over the risky one and determines the main comparative statics of these preferences given

[95] Note that in order to use equation (3.2), we think of any time lottery p as determining payoffs from next period on, preceded by a (known) consumption of c today.

the parameters of the model:

Proposition 3.3.1 *Under EZ, for any β, ρ, and x, there exists $\bar{\alpha}_{\rho,\beta,x} > \max\{\rho, 1\}$ such that $V_{EZ}\left(\delta_{(x,t)}\right) > V_{EZ}\left(\langle 0.5, (x, t-1); 0.5, (x, t+1)\rangle\right)$ if and only if $\alpha > \bar{\alpha}_{\rho,\beta,x}$. Moreover, $\lim_{x \searrow 0} \bar{\alpha}_{\rho,\beta,x} = +\infty$.*

Proof See Appendix.

Proposition 3.3.1 shows that, controlling for discounting β, elasticity of intertemporal substitution $1/\rho$, and the size of the prize x, more risk averse individuals are more likely to prefer the safe lottery.[96] That is, under EZ there is a connection between risk aversion over time lotteries and risk aversion over regular, atemporal lotteries. Moreover, the *risky* lottery is always preferred if the utility function is less concave than a logarithmic function ($\alpha < 1$), if $\alpha \leq \rho$, or if the prize is small enough.

To illustrate the nature of this trade-off, and why EZ allows for a preference for the safe lottery while DEU does not, consider the case of an infinite elasticity of intertemporal substitution ($\rho = 0$) and suppose the mean payment period is $t = 3$. Applying equation (3.3), the value of the safe lottery is simply the (per-period) discounted present value of consumption:

$$(1-\beta)\, c \left[1 + \beta + \lambda\beta^2 + \frac{\beta^3}{1-\beta}\right], \tag{3.5}$$

The value when choosing the risky lottery (the $50:50$ mixture between $t = 2$ and $t = 4$) equals

$$(1-\beta)\, c \left[1 + \beta \frac{\left(\lambda + \frac{\beta}{1-\beta}\right)^{1-\alpha} + \left(1 + \beta + \lambda\beta^2 + \frac{\beta^3}{1-\beta}\right)^{1-\alpha}}{2}\right]^{\frac{1}{1-\alpha}}, \tag{3.6}$$

which is the (per-period) certainty equivalent of the atemporal lottery that pays the discounted present value of future payments, that is, it pays either $c\left(\lambda + \frac{\beta}{1-\beta}\right)$, or $c\left(1 + \beta + \lambda\beta^2 + \frac{\beta^3}{1-\beta}\right)$ with equal probabilities.

[96] That is, $\alpha' > \alpha$ implies that if the decision maker with coefficient of risk aversion α prefers the safe lottery over the risky one, so does the decision maker with α' (holding other parameters fixed).

From (3.6), it follows that the value of the risky lottery is *decreasing* in the risk aversion α, while that of the safe lottery is unaffected. With risk neutrality ($\alpha = 0$), the risky lottery is preferred since it offers a higher expected discounted payment. As risk aversion α goes to infinity, however, the value of choosing the risky lottery decreases to that of receiving the worst possible outcome for sure – namely, the present discounted value from getting the prize at $t = 4$. Since receiving the prize at $t = 3$ is better than receiving it at $t = 4$, the safe lottery is preferred with extreme risk aversion. By monotonicity and continuity, there exists a unique cutoff $\bar{\alpha}_{0,\beta,x}$ separating the regions where safe lottery and the risky lottery are preferred. Since EZ coincides with DEU when $\alpha = \rho$, and DEU displays RSTL, the cutoff level of risk aversion $\bar{\alpha}_{\rho,\beta,x}$ must be greater than ρ.[97] Further notice that the cutoff $\bar{\alpha}_{\rho,\beta,x}$ only depends on the ratio x/c and does not depend on the time distance t. This follows from the homogeneity of this version of EZ and the dynamic consistency property of these preferences.

A broader intuition of why EZ preferences need not be RSTL is the following. Under DEU, the decision maker evaluates the risky lottery by first computing the discounted utility of *each* possible consumption path, and then aggregating these values linearly, by taking expectation; no curvature, or distortion, is applied to this aggregation – in DEU all the curvature is applied when computing the value of each path. Instead, under EZ the decision maker also computes the discounted utility of each path, but then aggregate them *non-linearly*: she calculates the certainty equivalent of the lottery over them in a way that depends on the individual's risk aversion (as captured by the parameter α). If risk aversion is sufficiently high, then she will prefer the safe lottery.

At the same time, Proposition 3.3.1 also shows that $\lim_{x \searrow 0} \bar{\alpha}_{\rho,\beta,x} = +\infty$, which means that for any α, ρ and β, EZ preferences are locally RSTL at x whenever x is small enough.

[97] Starting with Kreps and Porteus (1978), a large literature has studied preferences over the timing of resolution of uncertainty. With EZ, early resolution of uncertainty is preferred if and only if $\alpha > \rho$ (Epstein et al., 2014). Proposition 3.3.1 then implies that this condition is also needed for the safe time lottery to be preferred.

Therefore, with this formulation of EZ, preferences also cannot be locally RATL at any t (see footnote 93). Intuitively, this follows from the fact that the certainty equivalent of future paths is calculated according to Expected Utility (with constant relative risk aversion). Recall that Expected Utility implies that preferences are approximately risk neutral when stakes are very small (Segal and Spivak, 1990; Rabin, 2000). Then, since risk neutral preferences are RSTL, the risky lottery is preferred if the prize x is small enough.

We conclude with a brief discussion on the parameter restrictions on EZ implied by Proposition 3.3.1. Figure 11 presents the loci of points separating the regions where the safe and risky lotteries are preferred for a discount parameter of $\beta = 0.9$. Points above each curve correspond to parameters that favor the safe lottery; below each curve, the risky lottery is preferred. Notice that the region where the safe lottery is preferred increases with the prize x and with the elasticity of intertemporal substitution (i.e., it decreases with ρ). For example, with $x = c$, $\rho = 0.6$, and $\beta = 0.9$, the safe lottery is chosen as long as α is at least 15.

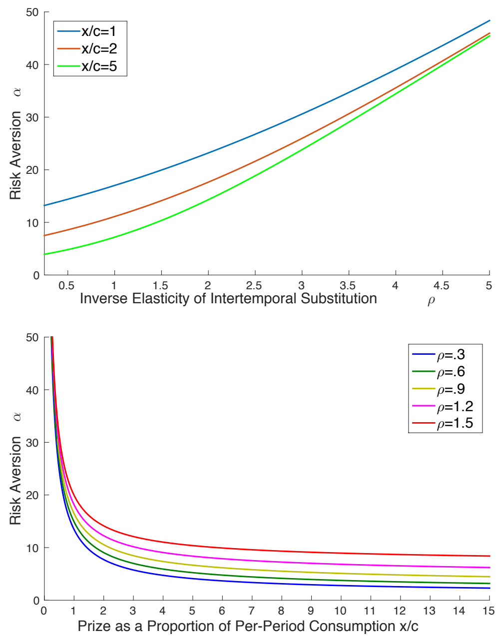
3.3.2. Discounted Probability Weighting Utility

An alternative way to generalize DEU is to allow for probability distortions. Consider a decision maker who evaluates a time lottery that pays prize $x > 0$ at t_0 with probability α and at $t_1 \geqslant t_0$ with probability $1 - \alpha$ according to

$$V_{DPWU}\left(\langle \alpha, (x, t_0); (1-\alpha), (x, t_1) \rangle\right) = \pi(\alpha) V_{DEU}\left(\delta_{(x,t_0)}\right) + (1 - \pi(\alpha)) V_{DEU}\left(\delta_{(x,t_1)}\right),$$

where $\pi : [0, 1] \to [0, 1]$ is increasing and continuous. This formula generalizes many non-Expected Utility models such as rank dependent utility (Quiggin, 1982; Yaari, 1987), cumulative prospect theory (Tversky and Kahneman, 1992), and disappointment aversion

Figure 11: The function $\bar{\alpha}_{\rho,\beta,x}$ determines whether the risky lottery is preferred to the safe lottery



Notes. Each panel depicts parameters under which a decision maker with discount parameter $\beta = 0.9$ will be indifferent between the safe lottery $\delta_{(x,3)}$ and the risky lottery $\langle 0.5, (x, 2); 0.5, (x, 4)\rangle$, where points above each curve correspond to parameters for which the safe lottery is preferred. Panel (a) has ρ on the horizontal axis and each curve corresponds to the prize x as a proportion of per-period consumption c. Panel (b) has the prize as a proportion of per-period consumption in the horizontal axis. Notice that the safe lottery is preferred when α, ρ, and x/c are large. As x/c decreases to zero, the decision maker eventually prefers the risky lottery.

(Gul, 1991).[98] For concreteness, we refer to any preferences satisfying this condition as Discounted Probability Weighting Utility (DPWU).

Consider again a choice between the safe lottery $\delta_{(x,t)}$ and the risky lottery that pays either $(x, t-1)$ or $(x, t+1)$ with equal probabilities. We now show that within DPWU, the more convex the discount function is, the more likely that the risky lottery is preferred; and the more probabilities are underweighted (that is, the lower $\pi(\alpha)$ is for any α), the more likely that the safe lottery is preferred. To make this statement precise, we need a measure of the convexity of the discount function. Let

$$d_t \equiv \frac{D(t-1) - D(t)}{D(t) - D(t+1)}$$

denote the degree of convexity at t. In particular with exponential discounting we have $d_t = \frac{1}{\beta}$, so a lower discount parameter implies a more convex discount function.

Straightforward calculations show that the safe time lottery is preferred if and only if

$$D(t) \geq \pi(0.5) D(t-1) + (1 - \pi(0.5)) D(t+1),$$

which can be rearranged as

$$d_t \leq \frac{1}{\pi(0.5)} - 1. \tag{3.7}$$

Thus, the safe time lottery is chosen by individuals with less convex discount functions and those who underweight probability more, showing that DPWU can accommodate local RATL at any t. Moreover, unlike with EZ, the condition for the safe lottery to be chosen (3.7) does not depend on the prize x.

For example, with exponential discounting ($D(t) = \beta^t$), the decision maker prefers the safe lottery if and only if $\beta \in \left(\sqrt{\frac{\pi(0.5)}{1-\pi(0.5)}}, 1\right)$. Thus, if the decision maker distorts probabilities

[98] In the context of static choice under risk and uncertainty, Ghirardato and Marinacci (2001) studied the general class of preferences with this property, which they termed biseparable.

pessimistically – leading to $\pi(0.5) < 0.5$ and thus $\frac{\pi(0.5)}{1-\pi(0.5)} < 1$ – and if β is high enough – the discount function is not too convex – then she will choose the safe lottery.

The following proposition summarizes the result from above and establishes that DPWU cannot accommodate preferences that are locally risk averse at any x (and, therefore, cannot be globally RATL).

Proposition 3.3.2 *Under DPWU, preferences are locally RATL at time t if and only if (3.7) holds. Moreover, they cannot be locally RATL at any prize x.*

Proof See Appendix.

3.4. Experiment

3.4.1. Design

We conducted the experiment at the Wharton Behavioral Lab at the Wharton School of the University of Pennsylvania in October and November 2013. We ran 16 sessions with 197 subjects, recruited from a database of volunteer students. Each session lasted about 45 minutes and no subject participated in more than one session. The experiment was conducted on paper-and-pencil questionnaire. No time limit was given to answer each question. Average earnings were \$25.20, including a \$10 show-up fee. Appendix D includes the complete questionnaire and instructions.

Some of the questions in the experiment involved payments to be made on the same day. If selected for payment, those would be paid at the end of each session, along with the show-up fee. Other questions involved payments to be made in the future. For these, subjects were told that their payment would be available to pick up from the lab starting from the date indicated. All payment dates were expressed in weeks, with the goal of reducing heterogeneity in transaction costs between the dates, under the assumption that students have a regular schedule each week during the semester.[99]

We ran two treatments: 'long delay' and 'short delay' (labeled Long and Short). A total of 91 and 105 subjects participated in each, respectively. Both treatments had identical questionnaires except for the length of delays in some of the questions: in the Long treatment, some payments were delayed by up to 12 weeks; in the Short treatment, the maximum delay was 5 weeks. Testing both treatments allows us to study long times spans, where differences between time lotteries become more pronounced; as well as shorter ones, where students' schedules are more stable, reducing heterogeneous sources of variation.[100]

The experiment has three parts.[101] Subjects received general instructions about the experiment and specific instructions about the first part of the experiment when they entered the room. Separate instructions were distributed in the beginning of each of the following parts.

Part I asks subjects to choose between different time lotteries and is the key part of our experiment. For example, the first question asked them to choose between \$15 in 2 weeks or \$15 in 1 week with probability .75 and in 5 weeks with probability .25. Subjects answered five questions of this kind. Table 23 lists the questions asked in each treatment. All questions offered two options that paid the same prize at different dates, and one distribution of

[99] An email was then sent to remind them of the approaching date (and they were told they would receive it). Subjects were also given the contact details of one of the authors, Daniel Gottlieb (at the time a full-time faculty at Wharton), in case they had questions about payments. Returning to the lab to collect the payment involve transaction costs, a typical concern for experiments involving delayed payments. However, transaction costs are less concerning for our experiment where *all* payments related to time lotteries take place in future dates. A second concern is that our questions involve payments of different amounts of money over time, and it is well-known that using money to study time-preferences may be problematic (Augenblick et al., forthcoming). However, for the purpose of our study most of these concerns do not apply (for example, the curvature of the utility function is inconsequential on ranking of time lotteries, as we have seen). Moreover, an important part of our analysis is the relationship between risk aversion over time lotteries and atemporal risk aversion. Since atemporal risk aversion is only defined for monetary lotteries, we focus our experiment on lotteries over money and leave for future research an investigation of time lotteries involving different objects.

[100] In the Short version all payments were scheduled before the end of the semester in which the experiment was run, and no payment was scheduled during exam week.

[101] The order of parts and of questions in each part were partly randomized: we first describe each part, and then discuss the randomization procedure.

payment dates was a mean preserving spread of the other. Hence, these questions allow us to elicit subjects' attitudes towards time lotteries. In three of them, one of the options had a known date; in the other two, both options had random payment dates. All subjects received the same first question (Question 1 in Table 23) in a separate sheet of paper. The answer to this question is a key indication of the subjects' preferences, as it captures their immediate reaction to this choice, uncontaminated by other questions.[102]

Table 23: Questions in Part I

	Long Delay				Short Delay			
Q.	$$	Option 1	*vs.*	Option 2	$$	Option 1	*vs.*	Option 2
1	$20	2 wk		75% 1 wk, 25% 5 wk	$20	2 wk		75% 1 wk, 25% 5 wk
2	$15	3 wk		90% 2 wk, 10% 12 wk	$15	3 wk		50% 1 wk, 50% 5 wk
3	$10	2 wk		50% 1 wk, 50% 3 wk	$10	2 wk		50% 1 wk, 50% 3 wk
4	$20	50% 2 wk, 50% 3 wk		50% 1 wk, 50% 4 wk	$20	50% 2 wk, 50% 3 wk		75% 2 wk, 25% 4 wk
5	$15	50% 2 wk, 50% 5 wk		75% 1 wk, 25% 11 wk	$10	50% 2 wk, 50% 5 wk		75% 3 wk, 25% 5 wk

Notes. Each lottery pays the same prize with different delays (in weeks). Subjects in the long delay treatment chose between 'Option 1' and 'Option 2, Long Delay.' Those in the short delay treatment chose between 'Option 1' and 'Option 2, Short Delay.'

Parts II and III use the multiple price list (MPL) method of Holt and Laury (2002) to measure time and risk preferences. With MPL, each question has a table with two columns and twenty-one rows. One column always displays the same option – e.g., receiving $10 today – while each row in the right column offers a slightly better option in constant increments. For example, in the first question, the options on the right go from $10 in 2 weeks to $15 in 2 weeks with .25$c$ increments per row. For each row, subjects have to choose between the left and the right options. These questions are typically interpreted as follows: if a subject chooses the option on the left for all rows above a point, and the option on the right below that point, then the indifference point should be where the switching takes place. Subjects who understand the procedure should not switch more than once.

Part II measures time preferences and attitudes towards time lotteries (see Table 24). Ques-

[102]One potential concern with offering a list of similar questions is that subjects may 'try' different answers even if they have a mild preference in one direction with some hedging concern in mind (Agranov and Ortoleva, 2015).

tions 6-9 measure discounting between various dates, allowing us also to quantify its convexity. Questions 10 and 11 allow us to quantify risk preferences towards time lotteries. Part III measures atemporal risk preferences, with payments taking place immediately at the end of the session (see Table 25). From these questions, we can quantify each subject's (standard) risk aversion. Moreover, by asking Allais' common-ratio-type questions (see for example Questions 12 and 13), we can determine which subjects behave according to Expected Utility theory and quantify violations of it. Finally, at the end of the experiment subjects answered a non-incentivized questionnaire.

Table 24: Questions in Part II

	Long Delay			Short Delay		
Q.	Option 1	*vs.*	Option 2	Option 1	*vs.*	Option 2
6	\$10 today		x in 2 wk	\$10 today		x in 2 wk
7	\$10 in 1 wk		x in 2 wk	\$10 in 1 wk		x in 2 wk
8	\$10 in 1 wk		x in 5 wk	\$10 in 1 wk		x in 3 wk
9	\$10 in 1 wk		x in 12 wk	\$10 in 1 wk		x in 4 wk
10	\$20 in 4 wk		\$20, x% in 2wk, (1-x)% in 12wk	\$25 in 3 wk		\$25, x% in 2wk, (1-x)% in 5wk
11	\$25 in 2 wk		\$25, x% in 1wk, (1-x)% in 5wk	\$25 in 2 wk		\$25, x% in 1wk, (1-x)% in 5wk

Notes. Questions 6-9 ask the amount \$$x$ that would make subjects indifferent between each option. Questions 10-11 ask the probability x% that would make subjects indifferent between each option. These amounts were determined using MPL.

Table 25: Questions in Part III

Q.	Option 1	*vs.*	Option 2
12	\$15		x% of \$20, (1-x)% of \$8
13	50% of \$15, 50% of \$8		x% of \$20, (1-x)% of \$8
14	20% of \$15, 80% of \$8		x% of \$20, (1-x)% of \$8
15	\$20		x% of \$30, (1-x)% of \$5
16	50% of \$20, 50% of \$5		x% of \$30, (1-x)% of \$3
17	10% of \$20, 90% of \$5		x% of \$30, (1-x)% of \$3

Notes. Questions ask the probability x% that would make subjects indifferent between each option, determined using MPL. All payments were scheduled for the day of the experiment.

After all subjects completed the questionnaire, one question was randomly selected from Parts I, II, and III for payment. The randomization of the question selected for payment, as well as the outcome of any lottery (if the selected question had random payments), was made with the use of dice.[103] Crucially, all uncertainty was resolved at the end of the experiment, including the one regarding payment dates. The instructions explicitly stated that subjects would know all payment dates before leaving the room.

The order of parts and of questions within parts was partly randomized. All randomizations took place at a session level. Therefore, all subjects in the same session saw the same questionnaire, but different questionnaires were used in different sessions. Because Part I is the key part of the experiment, all subjects saw it first to avoid contamination. For the same reason, within Part I, Question 1 was always the same. The other elements were randomized.[104] We find no significant effects of ordering.[105]

We conclude with a short discussion of our incentive scheme. The random payment mechanism, as well as the multiple price list method, are incentive compatible for Expected Utility maximizers, but not necessarily for more general preferences over risk.[106] Since this is the procedure used by most studies, a significant methodological work has been done to examine whether this creates relevant differences, with some reassuring results.[107]

[103]Specifically, at the end of the experiment one participant was selected as 'the assistant,' using the roll of a die by the experimenter. This subject was then in charge of rolling the die and checking the outcome to determine payments. This was done to reduce the fear that the experimenter could manipulate the outcome. All was clearly explained in the initial instructions.

[104]Specifically: for questions in Part I other than the first, half of the subjects answered questions in one specific order (the one used above), while the other half used a randomized order. In each of them, which option would appear on the left and which on the right was also determined randomly. The order of Parts II and III was randomized. For both parts, it was determined randomly whether in the MPL the constant option would appear on the left or on the right. This was done (independently) for each part, but not for each question within a part: in Part II or III the constant option of the MPL was either on the left or on the right for all questions of that part. This is typical for experiments that use the MPL method, as it makes the procedure easier to explain.

[105]The only exception is that out of the five questions in the first part, subjects have a significant (moderate) preference for the option on the right in the second question. While this is most likely a spurious significance (due to the large number of tests run), the order was randomized for all sessions and thus this tendency should have no impact on our analysis.

[106]Holt (1986) points out that a subject who obeys the reduction of compound lotteries but violates the independence axiom may make different choices under a randomly incentivized elicitation procedure than he would make in each choice in isolation. Conversely, if the decision maker treats compound lotteries by first assessing the certainty equivalents of all first stage lotteries and then plugging these numbers into a second stage lottery (as in Segal, 1990), then this procedure is incentive compatible. Karni and Safra (1987) prove the non-existence of an incentive compatible mechanism for general non-Expected Utility preferences.

[107]Beattie and Loomes (1997), Cubitt et al. (1998) and Hey and Lee (2005) all compare the behavior of subjects in randomly incentivized treatments to those that answer just one choice, and find little difference. Also encouragingly, Kurata et al. (2009) compare the behavior of subjects that do and do not violate

3.4.2. Results

We start with two preliminary results. First, most subjects (82%) exhibit convex discounting. Second, the large majority of subjects gave monotone answers to MPL questions: 13% gave a non-monotone answer in at least one of the 12 MPL questions, and only 4.6% gave non-monotone answers in more than one. These are substantially lower numbers (i.e., fewer violations) than what previous studies have found (Holt and Laury, 2002).[108]

Risk Aversion over Time Lotteries

Our main variable of interest is each subject's risk attitude towards time lotteries, which can be measured in three different ways. First, we can measure it using Question 1 of Part I, the first question that subjects see. Second, we can look at the answers to all five questions in Part I and ask whether subjects exhibited RATL in the majority of them (for the purpose of this section, we say that subjects are RATL in a given question if, in that question, they chose the option with the smallest variance of the payment date). A third way is to look at the answers given in Questions 10 and 11 of Part II, using MPL.

Table 26 presents the percentage of RATL answers for each of these measures. The results are consistent: in most questions, especially in the Long treatment, the majority of subjects are RATL. In Question 1, the proportions are about 66% for the Long treatment and 56% for the Short one. Notice that subjects are still RATL when *both* options are risky but one of the options is a mean preserving spread of the other (Questions 4 and 5). Thus, the data suggest an aversion to mean preserving spreads, not simply an attraction towards certainty.[109] In addition, the majority of subjects are RATL in at least one of the two MPL

Expected Utility in the Becker-DeGroot-Marschak procedure (which is strategically equivalent to MPL) and find no difference. On the other hand, Freeman et al. (2015) find that subjects tend to choose the riskier lottery more often in choices from lists than in pairwise choices.

[108]The non-monotone behavior did not concentrate in any specific question. Following the typical approach in the literature, these answers are disregarded. Alternatively, we could have dropped any subject that exhibits a non-monotone behavior at least once. Doing so does leave our results essentially unchanged (as to be expected given the number of them).

[109]We also note that we find no gender differences in RATL, or relationship with results in the SATMath

questions (Questions 10 and 11).

Table 26: Percentage of RATL in each question

Question	Long	Short
1	65.71	56.04
2	50.48	54.95
3	48.57	37.36
4	64.76	38.46
5	73.33	52.75
Majority in 1-5	64.76	49.45
10	44.23	54.44
11	57.28	41.11
Either in 10 or 11	64.07	66.66

Table 27: Frequency of RATL answers in Part I

Frequency of RATL	Long Delay		Short Delay	
	Percent	Cum.	Percent	Cum.
0	2.86	2.86	9.89	9.89
1	9.52	12.38	16.48	26.37
2	22.86	35.24	24.28	50.55
3	23.81	59.05	26.37	76.92
4	28.57	87.62	19.78	96.70
5	12.38	100.00	3.30	100

In most questions, RATL is stronger in the Long rather than in the Short treatment.[110] This is intuitive: when the time horizon is relatively short, the difference between the options decreases and subjects should become closer to being indifferent – and their choices closer to an even split. In the Long treatment the difference in time horizon increases, and so does the differences between the options. While the standard model suggests that this should push more strongly towards RSTL, the *opposite* holds in our data.

While most answers are consistent with RATL, it could be that a non-trivial fraction of our subjects still consistently chooses the risky option, as predicted by DEU. However, Table 27 shows that the fraction of subjects who does so is minuscule in the Long treatment (2.86%) and very small in the Short one (9.89 %). By contrast, in the Long treatment almost 41% give risk averse answers at least 4 out of 5 times, and 59% at least three times. (These numbers are about 23% and 48.45% in the short treatment.)

Result 3.4.1 *The majority of subjects is RATL. This is more pronounced in the Long*

or the total SAT.

[110]In Question 10, subjects appear to be more RSTL in the Long than in the Short treatment. This could reflect a genuine preference, or it could be because of the specifics of the MPL for this question: to be RATL in the Long treatment, a subject would have to 'switch' very close to the end of the list (row 18 out of 21). It is well-known that with MPL subjects tend to switch close to the middle of the table, generating a bias towards RSTL in this case. (In all other cases, the switching point exhibited by an RATL subject was after the middle but closer to it.)

rather than Short treatment. Only a very small fraction is consistently RSTL.

RATL and convexity, probability weighting, and atemporal risk aversion

In this subsection we analyze the relationship between RATL and convex discounting, violations of Expected Utility, and atemporal risk aversion. With DEU, all subjects with convex discounting should be RSTL; in turn, this means that such tendency should be negatively related to convexity of the discount function. With EZ, RATL should be positively correlated with atemporal risk aversion. With DPWU, only subjects who underweight probabilities should exhibit RATL.

We test these predictions using our data. We quantify convexity of the discount function, violations of Expected Utility, and atemporal risk aversion using the MPL measures collected in Parts II and III. We first analyze RATL in the restricted samples of subjects with convex discounting, those who behave according to Expected Utility, and those who do not underweight probabilities. We then move to the regression analysis.

As previously described, we determine which subjects have convex discounting based on their answers to Questions 7, 8, and 9 in Part II.

There are two related measures of violations of Expected Utility, constructed using questions from Part III. The first measure uses answers from Questions 12 and 13, or 12 and 14 (see Table 25) to determine if subjects display what is typically called certainty bias (Kahneman and Tversky, 1979).[111] This is the relevant measure for our analysis since subjects who

[111] Suppose that in Question 12 the subject switches at x_{12}, while in Question 13 she switches at x_{13}. If the subject follows Expected Utility, we should have $2x_{13} = x_{12}$. A certainty-biased subject would instead have $x_{12} > 2x_{13}$; because she is attracted by the certainty of Option 1 in Question 12, she demands a high probability of receiving the high prize in Option 2 to be indifferent. Thus, the answers to Question 12 and 13 allow us to identify subjects who are certainty biased and to quantify the bias (by $x_{12} - 2x_{13}$). A similar measure can be obtained from the answers to Questions 12 and 14. In what follows, when we want to focus on subjects who are certainty biased, we focus on the measure obtained from Questions 12 and 13 (the results using the measure obtained from Questions 12 and 14 are essentially identical and are reported in Appendix). When we need to quantify certainty bias (in the regression analysis), we use instead the principal component of the two measures, which should reduce the observation error (essentially identical results hold using either of the two measures or their average.)

underweight probabilities must also display certainty bias. We find that a small number of subjects do (15.71%).[112]

The second measure uses the answers to Questions 12, 13, and 14 to determine whether they are jointly consistent with Expected Utility. Since this is a very demanding requirement (it is well-known that these measures can be noisy), we consider as "approximately Expected Utility" those subjects who abide by Expected Utility in all three questions, allowing for a "one-line mistake" – more precisely, they would be consistent with Expected Utility if we changed their answer to these questions by one line. These are 39.89% of the pool (43.3% and 36.67% in the Long and Short treatments, respectively).

Table 28: Proportion of RATL subjects

Sample	Convex Discounting		Approximately Expected Utility		No Certainty Bias	
Treatment	Long	Short	Long	Short	Long	Short
Question 1	67.78	50.70*	66.67	60.61	67.50	55.00
Majority in Q1-5	65.56	43.66**	64.29	60.61*	68.75*	50.00
Question 10	46.07	52.86	54.76*	68.75**	47.50	51.90
Question 11	57.95	44.29**	64.29	56.25	54.43	48.10
Observations	90	71	42	33	80	80

Notes. The first row measures RATL using Question 1. The second row identifies as RATL subjects who chose the safe option in the majority of Questions 1-5. The third and fourth rows use answers to MPL Questions 10 and 11. Columns present the proportion of RATL subjects in the subsamples of subjects with convex discounting, approximately Expected Utility, and those with no Certainty Bias as measured using Questions 12 and 13. * and ** denote significance at the 10% and 5% level in a Chi-squared test of whether each subset is different from its complement.

Table 28 shows that, based on the four different measures, subjects are still RATL in each of the subsamples described above. The table also shows the results of Chi-squared tests on whether subjects in each of subsample are statistically different from those outside of it. Using Question 1, subjects in any of these subsamples are not statistically different from those in their complements at the 5% level. Moreover, in all treatments and for all RATL measures, certainty-biased subjects are statistically indistinguishable from those

[112]These small numbers are not surprising: it is a stylized fact that certainty bias is less frequent when stakes are small, as in our experiment (Conlisk, 1989; Camerer, 1989; Burke et al., 1996; Fan, 2002; Huck and Müller, 2012). See the discussion in Cerreia-Vioglio et al. (2015).

with no certainty bias at the 5% level. Similarly, approximately Expected Utility subjects were statistically different from the rest of the population only for the Short treatment and only for Question 10 – where approximately Expected Utility subjects tend to be *more* RATL (the opposite of the prediction of DPWU). Subjects with convex discounting were only statistically different for the Short treatments in the Majority of Questions 1-5 and in Question 11.[113] These results are in direct contrast with the predictions of DEU and probability weighting models: according to the former, there should be no RATL with convex discounting; according to the latter, there should be no RATL without certainty bias, or within approximately Expected Utility subjects.

A regression analysis confirms these results. Table 29 presents the coefficients of Probit regressions of RATL with our measures of certainty bias and of convexity as independent variables. (Essentially identical results hold when considering each in isolation; see Appendix.) Two patterns emerge. First, certainty bias is generally not related to RATL. With the exception of the Short treatment in Question 10 (regression 6), certainty bias is either statistically insignificant, even at the 10% level, or its sign is the opposite of what theory predicts. Second, there is only a significant relationship between convexity and RATL for the Short treatment, and only for the RATL measures using the majority of Questions 1-5 or Question 11 (regressions 4 and 8). In all other regressions, convexity is either insignificant at the 10% level, or it has the opposite sign relative to what theory predicts. Both patters contradict the predictions of DEU and DPWU.

Lastly, we examine the relationship between RATL and atemporal risk aversion. Table 30 presents the coefficients from a Probit regression, with our four RATL measures as dependent variables and the degree of risk aversion (as measured in Question 12) as the independent variable. Consistently with EZ, the coefficients are positive and, with the exception of the Short treatment in Question 1, they were all statistically significant at the

[113] Even this should be taken with caution: since we are simultaneously running 24 tests, this can easily be spurious.

Table 29: Probit Regressions: RATL and Convexity and Certainty Bias

Dep. Variable	RATL Q1		RATL Majority Q1-5		RATL Q10		RATL Q11	
Treatment	Long	Short	Long	Short	Long	Short	Long	Short
(Probit)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Certainty Bias	-.25*	.18	-.20	.18	-.03	.44***	.16	.23
	(-1.94)	(1.17)	(-1.60)	(1.20)	(-.25)	(2.77)	(1.42)	(1.52)
Convexity	4.27*	-4.45	.06	-11.10***	3.47	-1.77	.99	-7.73**
	(1.82)	(-1.29)	(.03)	(-2.86)	(1.64)	(-1.29)	(.46)	(-2.16)
Constant	.19	.28*	.39**	.19	-.26	.16	.19	.12
	(1.07)	(1.82)	(2.15)	(1.20)	(-1.52)	(1.03)	(1.08)	(.79)
Pseudo-R^2	.06	.02	.02	.01	.02	.07	.02	.06
Obs.	92	86	92	86	92	85	92	85

Notes. Dependent variables are indicated in the first row. Coefficients in brackets are z-statistics. *, **, and *** denote significance at the 10%, 5% and 1% level.

Table 30: Probit Regressions: RATL and Atemporal Risk Aversion

Dep. Variable	RATL Q.1		RATL Majority Q.1-5		RATL Q.10		RATL Q.11	
Treatment	Long	Short	Long	Short	Long	Short	Long	Short
(Probit)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk Aversion,	.336**	.175	.308**	.341**	.459***	.435***	.571***	.239***
Atemporal	(2.41)	(1.30)	(2.27)	(2.40)	(3.27)	(2.91)	(3.81)	(1.75)
Constant	.07	-.01	.07	-.34*	-.60***	-.27	-.37***	-.19
	(0.38)	(-0.07)	(0.38)	(-1.79)	(-3.01)	(-1.47)	(-1.87)	(-1.08)
Pseudo-R^2	0.047	0.014	0.040	0.049	0.083	0.076	0.121	0.025
Observations	101	90	101	90	101	89	100	89

Notes. Dependent variables are indicated in the first row. Atemporal risk aversion measure is obtained from Question 12. RATL measures were obtained from Question 1 (Regressions 1 and 2), having chosen the safe option in the majority of Questions 1-5 (Regressions 3 and 4), and MPL Questions 10 and 11 (Regressions 5-8). Coefficients in brackets are z-statistics. *, **, and *** denote significance at the 10%, 5% and 1% level.

5% level. In fact, all treatments in Questions 10 and 11 were statistically significant at the 1% level. Similar results hold constructing risk aversion from Question 15 or using a linear probability model.

Result 3.4.2 *Subjects who exhibit convex discounting, no certainty bias, or are approximately Expected Utility also have a tendency to be RATL. In fact, the proportions in these groups are almost identical to the one in the overall population. Regression analysis shows* *that RATL is unrelated to probability distortion and generally unrelated to convexity. It is, however, related to (atemporal) risk aversion.*

These findings are not compatible with Discounted Expected Uitlity or with Discounted Probability Weighting Utility. Instead, they are compatible with Epstein Zin.

APPENDIX

A.1. Appendix

A.1.1. Chapter 1 Appendix Tables

Appendix Table
1. Previous Effort Instrumental Variable: Impact on Effort – Replication with Sliders

Dependent Variable:	Specification				
Sliders Moved Correctly	(1)	(2)	(3)	(4)	(5)
Sliders Previous Period	0.35	0.57***	0.46**	0.42**	0.42**
(aka ρ parameter)	(0.39)	(0.19)	(0.18)	(0.21)	(0.21)
Piece Rate	41.73***	48.27***	60.26***	58.21***	57.51***
(in cents per 100 sliders)	(11.80)	(12.17)	(13.39)	(13.51)	(13.72)
First Stage F Stat (IV)	5.3	18.9	20.2	16.6	15.1
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	52.1	52.1	52.1	52.1	52.1
Number of Observations	552	552	552	552	549
Number of Individuals	184	184	184	184	183
Adj-R^2	0.45	0.66	0.63	0.64	0.65

Notes. The above table represents an additional experiment that replicated the main findings using a different (slider) task. The dependent variable is the number of sliders correctly moved to 50% in a single period. See more details in Appendix section 10.3. All specifications report results from linear Instrumental Variable regressions estimated by (iterative) GMM and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 1 subject. The First Stage F Statistic is from the instrument alone (F(1, 183)). Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Appendix Table 2. Previous Effort Instrumental Variable: Impact on Effort – Experiment 1

Dependent Variable:	Specification					
Problems Solved	(1)	(2)	(3)	(4)	(5)	(6)
Problems Previous Period	0.49**	0.48*	−0.02	0.41	0.71*	0.47
	(0.25)	(0.26)	(0.17)	(0.34)	(0.37)	(0.40)
Piece Rate	4.50***	4.51***	2.03	7.41***	7.02***	6.13***
	(1.31)	(1.29)	(2.88)	(2.41)	(2.54)	(2.30)
Phone Access	−0.42**	−0.40**	0.23	0.14	0.09	0.17
	(0.19)	(0.18)	(0.27)	(0.34)	(0.34)	(0.33)
First Stage F Stat (IV)	9.8	9.5	4.6	2.8	3.6	2.8
PreTreatment Quintiles		X	X		X	X
Period Fixed Effects		X	X		X	X
Session Fixed Effects			X			X
Individual Controls			X			X
Period 1 and 2 Only				X	X	X
Dependent Variable Mean	7.85	7.85	7.85	7.79	7.79	7.79
Number of Observations	930	930	930	465	465	465
Number of Individuals	155	155	155	155	155	155
Adj-R^2	0.01	0.31	0.37	0.02	0.66	0.66

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from linear Instrumental Variable regressions estimated by (iterative) GMM and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 2 subjects. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Appendix Table 3. Previous Period Piece Rate and Phone Access: Random Effects – Experiment 2

Dependent Variable:			Specification		
Problems Solved	(1)	(2)	(3)	(4)	(5)
Piece Rate	11.05***	12.58***	14.95***	15.07***	15.00***
	(2.40)	(2.29)	(2.67)	(2.71)	(2.78)
Previous Period's Piece Rate	2.47	4.00	5.62**	5.73**	5.78**
	(2.68)	(2.55)	(2.80)	(2.83)	(2.89)
Phone Access	−0.05	−0.08	0.25	0.27	0.28
	(0.22)	(0.22)	(0.27)	(0.27)	(0.27)
Previous Period Phone Access	−0.05	−0.08	0.15	0.17	0.17
	(0.27)	(0.27)	(0.31)	(0.31)	(0.31)
Random Effects	X	X	X	X	X
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.01	0.41	0.41	0.44	0.45

Notes. The dependent variable is the number of problems solved correctly in a single period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Appendix Table 4. Previous Period Piece Rate and Phone Access: Fixed Effects – Experiment 2

Dependent Variable: Problems Solved	Specification (1)	(2)	(3)
Piece Rate	12.66***	10.46***	13.49***
	(3.51)	(2.54)	(3.00)
Previous Period's Piece Rate	4.09	1.89	4.16
	(4.03)	(2.63)	(2.94)
Phone Access	−0.03	−0.05	0.36
	(0.36)	(0.22)	(0.28)
Previous Period Phone Access	−0.03	−0.05	0.26
	(0.41)	(0.26)	(0.31)
Subject Fixed Effects		X	X
Period Fixed Effects			X
Dependent Variable Mean	7.23	7.23	7.23
Number of Observations	1266	1266	1266
Number of Individuals	422	422	422
Adj-R^2	0.01	0.72	0.72

Notes. In the presence of momentum, these estimates are **severely biased** downward. They are presented here only in the spirit of openness but are not intended to be taken as accurate estimates. The dependent variable is the number of problems solved correctly in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Note that session fixed effects and individual controls cannot be estimated with subject fixed effects as these variables do not vary within individual. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

Appendix Table 5. Previous Period Piece Rate and Phone Access: Impact on YouTube Searches – Experiment 2

Dependent Variable:	Specification				
Youtube Searches	(1)	(2)	(3)	(4)	(5)
Piece Rate	−2.39***	−2.52***	−2.80***	−2.68***	−2.59***
(in cents)	(0.42)	(0.44)	(0.56)	(0.63)	(0.65)
Previous Period's Piece Rate	−1.01*	−1.15**	−0.94	−0.83	−0.72
	(0.56)	(0.55)	(0.58)	(0.66)	(0.68)
Phone Access	−0.02	−0.02	−0.06	−0.03	−0.04
	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)
Previous Period Phone Access	0.09	0.09	0.12	0.14	0.14
	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)
PreTreatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	0.24	0.24	0.24	0.24	0.24
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.01	0.02	0.02	0.04	0.07

Notes. The dependent variable is the number of YouTube videos searched in a period. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, sex, ethnicity, computer skill test, and total # of experimental sessions done at the lab, but could not be matched for 2 subjects. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

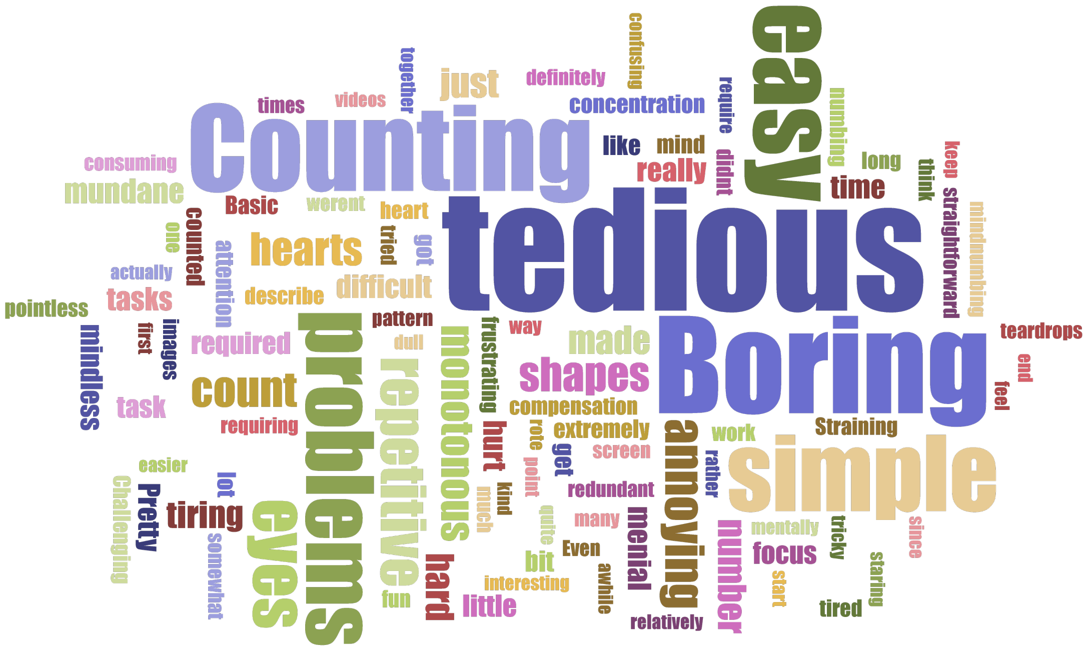
Appendix Table 6. Phone Usage: Relationship with Effort – Experiment 2

Dependent Variable:	Specification				
Problems Solved	(1)	(2)	(3)	(4)	(5)
Phone Usage	-3.81^{***}	-2.19^{***}	-2.06^{***}	-2.11^{***}	-1.93^{***}
	(0.88)	(0.71)	(0.72)	(0.73)	(0.73)
Previous Period Phone Usage	-4.73^{***}	-3.14^{***}	-3.11^{***}	-3.16^{***}	-2.99^{***}
	(0.92)	(0.65)	(0.67)	(0.69)	(0.70)
Piece Rate	12.41***	14.95***	15.99***	15.96***	16.13***
	(3.51)	(2.76)	(2.96)	(2.95)	(3.01)
Previous Period Piece Rate	3.95	6.50**	6.70**	6.67**	6.96**
	(3.93)	(3.05)	(3.23)	(3.24)	(3.29)
Pre-Treatment Quintiles		X	X	X	X
Period Fixed Effects			X	X	X
Session Fixed Effects				X	X
Individual Controls					X
Dependent Variable Mean	7.23	7.23	7.23	7.23	7.23
Number of Observations	1266	1266	1266	1266	1260
Number of Individuals	422	422	422	422	420
Adj-R^2	0.04	0.42	0.43	0.45	0.44

Notes. The dependent variable is the number of problems solved correctly in a period. Phone Usage is a self reported variable indicating use of the phone during period 3. As this is endogenously chosen, these regressions should not be taken as causal, as a subject who uses the phone may have unobservable differences. All specifications report results from OLS regressions and also include a constant term. PreTreatment Quintiles represent five binary variables to non-parametrically control for the number of problems subject solved in the pre-treatment training period. Individual Controls include age, ethnicity, computer skill test, and total # of experimental sessions done at the lab. Gender could not be matched for one subject, and the controls for an additional subject. Standard errors given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

A.1.2. Chapter 1 Appendix Figures

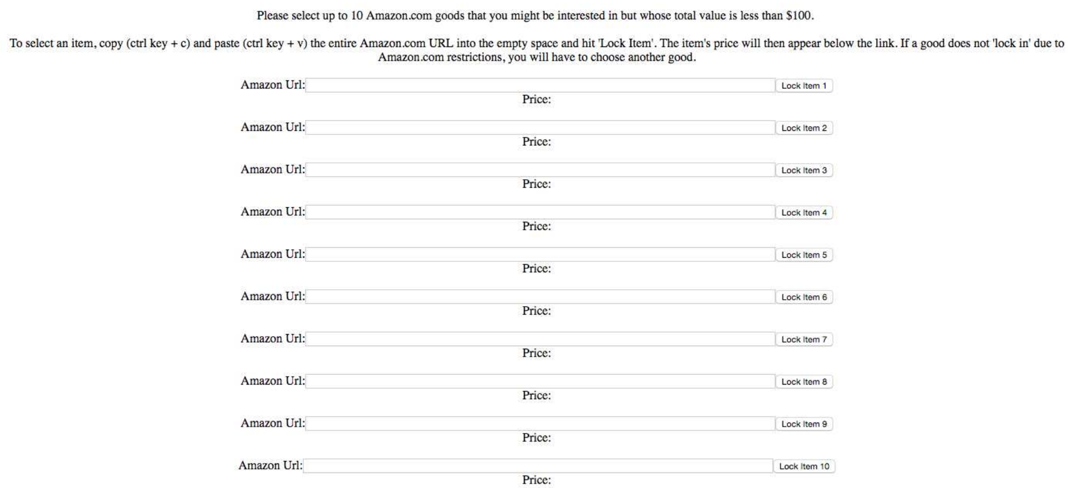
Appendix Figure 1. Word Cloud for Survey – "Opinion of Task"



Notes. Top 100 words from responses to a post experiment survey question asking "What is your opinion of the task?" Size scaled linearly with count.

Source: Jasondavies.com word cloud generator.

Appendix Figure 2. Quiz for Introduction Instructions


Please select up to 10 Amazon.com goods that you might be interested in but whose total value is less than $100.

To select an item, copy (ctrl key + c) and paste (ctrl key + v) the entire Amazon.com URL into the empty space and hit 'Lock Item'. The item's price will then appear below the link. If a good does not 'lock in' due to Amazon.com restrictions, you will have to choose another good.

Amazon Url: Lock Item 1
Price:

Amazon Url: Lock Item 2
Price:

Amazon Url: Lock Item 3
Price:

Amazon Url: Lock Item 4
Price:

Amazon Url: Lock Item 5
Price:

Amazon Url: Lock Item 6
Price:

Amazon Url: Lock Item 7
Price:

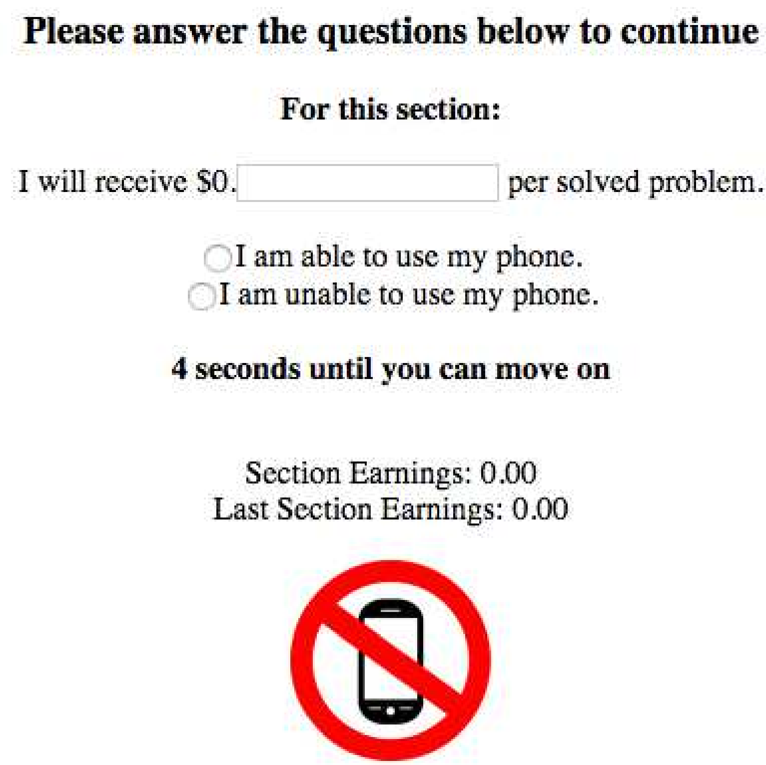
Amazon Url: Lock Item 8
Price:

Amazon Url: Lock Item 9
Price:

Amazon Url: Lock Item 10
Price:

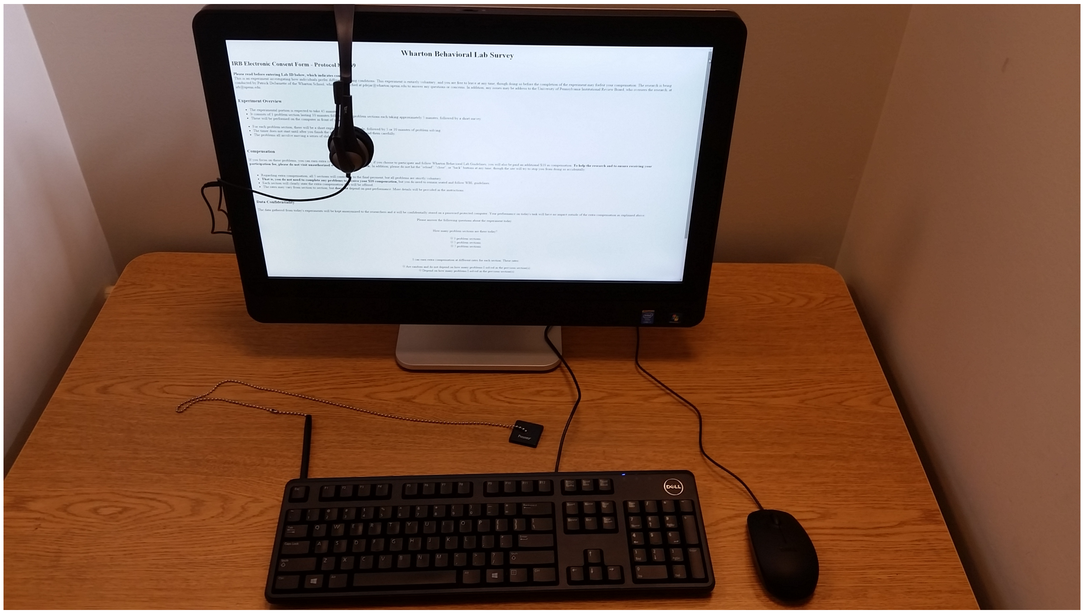
Notes. Every participant in experiment 2 had to answer the above questions after reading experiment instructions. Subjects had to answer all three questions correctly to proceed. If the subject entered the wrong answers, the browser would alert them to this and ask for them to review the instructions again.

Appendix Figure 3. Quiz for Instructions Prior to Each Period



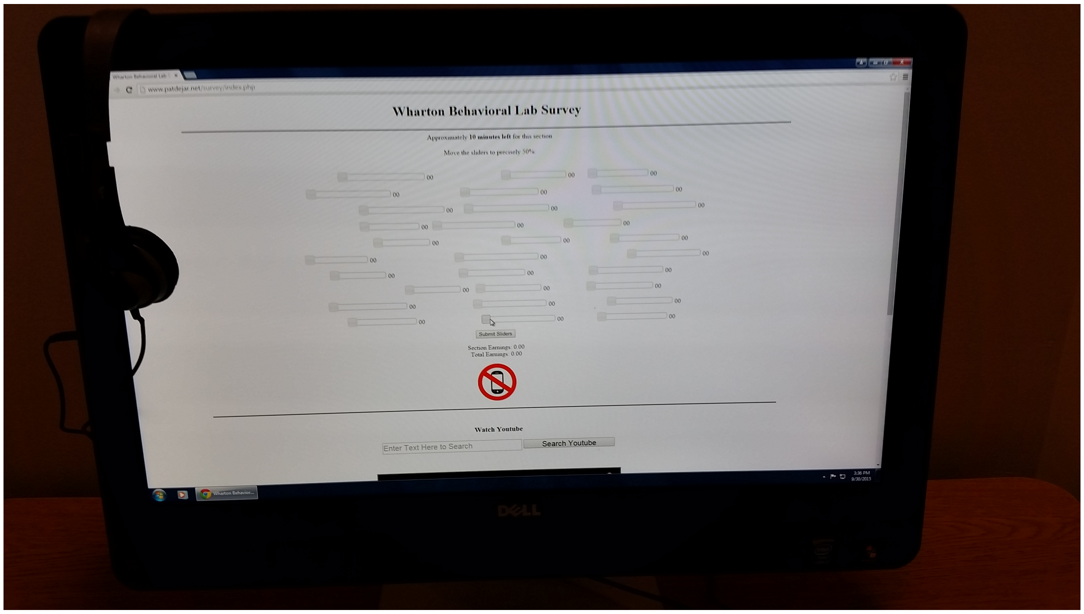
Notes. Every participant had to answer the following questions prior to every period (including Pre-Treatment). If the subject had information about future periods, they were also quizzed on the piece rate and phone access for future periods. If the subject entered the wrong answers, the browser would alert them to this and ask for them to review the instructions again.

Appendix Figure 4. Cubicle Environment



Notes. Every participant had access to an identical computer with headphones as pictured above. Screen brightness was uniformly set at 95% to ensure consistency across cubicles. It was not possible to see other subjects from within the cubicle. Google Chrome was employed as the browser during the task, with the window maximized (full screen mode). All instructions were written, but RAs were on site to answer any additional questions.

Appendix Figure 5. Slider Task Example



Notes. Slider Task employed as a replication of momentum effects. As can seen above, it would be difficult to view YouTube and move sliders at the same time on the monitor and resolution employed. Furthermore, the website code blocked any attempt to “zoom” in or out.

A.1.3. Chapter 1 – Slider Task (Experimental Replication)

In addition to replicating findings of the first experiment with Experiment 2 (details in main text), I also ran an additional experiment with a different task to serve as an additional replication. In this setting, subjects face 30 "sliders" on a screen, as in Gill and Prowse (2011). This can be seen in Appendix Figure 5 above. The subjects are asked to move the slider to exactly 50% of the way, with a numerical setting next to the slider indicating the current %. The exact position and length of sliders was randomized to make the task more difficult.

In this setting, as the sliders take far less time than the counting problems, the piece rate was also reduced. Subjects received a baseline of \$0.01 per 10 sliders. The "high piece rate" treatment was \$0.03 per 10 sliders (three times baseline), while the "low piece rate" treatment was \$0.0033 per 10 sliders (one third of baseline). Subjects were rounded to the nearest cent in the case they were unable to finish before time ran out. In addition, due to the small subject size available and the relatively "noisy" effect of cellphones in previous experiments, there was no phone access treatment. This affords us greater power in detecting effects through the financial incentives, but does make it harder to distinguish some theories ruled out in experiment 2.

In order to better understand the underlying model of momentum, the pre-treatment period was reduced 5 minutes and an additional 5 minute period following treatment was added. This allows one to investigate the rate of decay of effort inducement.

As predicted by a model of momentum, these gains from working harder continue to decay exponentially as time goes on. As the design and results are otherwise similar to the main findings of the paper, I have omitted most tables for brevity. See Appendix Table 1 or contact the author for additional details.

A.1.4. Chapter 1 Proofs

Proof of Proposition 1

This proof will be using methods of supermodularity discussed in Milgrom and Shannon (1995).[114] For simplicity, I retain the assumption about the utility function being twice differentiable over c_t, e_t, e_{t-1}, however this assumption could be weakened as long as the utility function maintains increasing differences.

MS Theorem 6a: I begin by applying Theorem 6 of Milgrom and Shannon (1995), which first states that

[114] An earlier draft of the proof uses the (Dini) Multivariate Implicit Function Theorem, but supermodularity allows for fewer restrictions and also removes the need for matrix manipulation in calculating determinants.

a twice differentiable function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ has increasing differences in (x, z) if and only if $\frac{\partial f^2}{\partial x_i \partial z_j} \geq 0$ for all $i = 1, ..., n$ and $j = 1, ..., m$.

MS Theorem 6a Conditions: In this case, $x_1 \equiv c_1$, $x_2 \equiv c_2$,..., $x_T \equiv c_T$, and $x_{T+1} \equiv e_1$, $x_{T+2} \equiv e_2, ..., x_{2T} \equiv e_T$. y are the non-choice state variables, $z_1 \equiv w_1$, $z_2 \equiv w_2$, ..., $z_T \equiv w_T$, and $z_{T+1} \equiv -\gamma_1$, $z_{T+1} \equiv -\gamma_2$, ...$z_{2T} \equiv -\gamma_T$, and f is the full Lagrangian (treating y_t, r, and p_t as fixed):

$$f(x, z) = \sum_{t=1}^{T} \delta^{t-1} u(c_t,\, e_t,\, e_{t-1},\, \gamma_t) + \lambda \left(\sum_{t=1}^{T} (w_t e_t + y_t)(1+r)^{-t} - \sum_{t=1}^{T} p_t c_t (1+r)^{-t} \right)$$

With variables redefined this way, we can check the conditions. For consumption and wage, it is clear that $\frac{\partial f^2}{\partial c_i \partial w_j} = 0$ as w_j does not enter the utility function and the budget constraint is linear with no term containing both c_i and w_j. For consumption and leisure technology, $\frac{\partial f^2}{\partial c_i \partial \gamma_j} = 0$ if $i \neq j$. However, without an assumption on $\frac{\partial f^2}{\partial c_t \partial \gamma_t}$, it is possible that increased leisure technology increases the utility of consumption so much that effort rises in every period (including time period t) to satisfy the greater demand for consumption goods. However, while leisure and consumption might have complementarities (and hence effort and consumption exhibit substitutibility), leisure technology itself is constructed to not influence the utility from consumption.[115]

For effort and wage, $\frac{\partial f^2}{\partial e_i \partial w_j} = 0$ if $i \neq j$ as wage and effort only appear together in the same time period. In this case $\frac{\partial f^2}{\partial e_t \partial w_t} = \frac{\partial f^2}{\partial e_t \partial w_t} = \lambda(1+r)^{-t}$, which is ≥ 0 as consumption c_t is enjoyable and $r > -1$.

For effort and leisure technology, $\frac{\partial f^2}{\partial e_i \partial \gamma_j} = 0$ if $i > j$ or $i + 1 < j$ as e_t only appears in two $u(\cdot)$ functions, at time t and time $t+1$. Thus, the only cases we need to check are on $\frac{\partial f^2}{\partial e_i \partial(-\gamma_i)}$ and $\frac{\partial f^2}{\partial e_i \partial(-\gamma_{i+1})}$, which are equivalent to $-\delta^{i-1} \frac{\partial u^2}{\partial e_i \partial \gamma_i}$ and $-\delta^{i} \frac{\partial u^2}{\partial e_i \partial \gamma_{i+1}}$, respectively. From the assumptions above, $\frac{\partial u^2}{\partial e_i \partial \gamma_i} \leq 0$, as leisure technology makes marginal contemporaneous effort more costly in utility terms, and $\frac{\partial u^2}{\partial e_i \partial \gamma_{i+1}} = 0$, as leisure technology does not carry over across periods. Thus, both $\frac{\partial f^2}{\partial e_i \partial(-\gamma_i)}$ and $\frac{\partial f^2}{\partial e_i \partial(-\gamma_{i+1})}$ are ≥ 0. And in conclusion, f has increasing differences in $(x,\, z)$.

MS Theorem 6b: I apply the second half of Theorem 6 of Milgrom and Shannon (1995), which states that a twice differentiable function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is supermodular in x if and only if $\frac{\partial f^2}{\partial x_i \partial x_j} \geq 0$ for all $i \neq j$ in $1, ..., n$.

[115]E.g. Popcorn is enjoyable while watching the new Star Wars movie; the new Star Wars movie does not make popcorn itself taste better on December 18th if you did not get to watch it.

MS Theorem 6b Conditions: As before, $x_1 \equiv c_1$, $x_2 \equiv c_2$,..., $x_T \equiv c_T$, and $x_{T+1} \equiv e_1$, $x_{T+2} \equiv e_2$, ..., $x_{2T} \equiv e_T$, and f is the full Lagrangian:

$$f(x,z) \quad = \quad \sum_{t=1}^{T} \delta^{t-1} u(c_t,\, e_t,\, e_{t-1},\, \gamma_t) + \lambda \left(\sum_{t=1}^{T} (w_t e_t + y_t)(1+r)^{-t} - \sum_{t=1}^{T} p_t c_t (1+r)^{-t} \right)$$

Note that $\frac{\partial f^2}{\partial c_i \partial c_j} = 0$ for $i \neq j$ as there is no overlap in the additively separable terms. Likewise $\frac{\partial f^2}{\partial e_i \partial e_j} = 0$ if $i > j$ or $i+1 < j$ as only last period's effort influences this period's effort.[116] Thus, for time period t we are concerned with two terms: first, $\frac{\partial f^2}{\partial e_t \partial e_{t-1}} = \delta^{t-1} \frac{\partial u^2}{\partial e_t \partial e_{t-1}} \geq 0$, as we have positive momentum (last period effort makes this period's effort marginally less costly in terms of utility). Second, $\frac{\partial f^2}{\partial e_{t+1} \partial e_t} = \delta^t \frac{\partial u^2}{\partial e_{t+1} \partial e_t} \geq 0$ for the same reasons.

The only remaining terms of interest are the cross-partials between consumption and effort. Similar to above, $\frac{\partial f^2}{\partial e_i \partial c_j} = 0$ if $i > j$ or $i+1 < j$ as the overlap only occurs for a c_t and e_t or c_t and e_{t-1}, as per the utility function. First, $\frac{\partial f^2}{\partial c_t \partial e_t} = \delta^{t-1} \frac{\partial u^2}{\partial c_t \partial e_t} \geq 0$, as assumed in the set up. This implies that consumption and effort are not complements. By a similar assumption $\frac{\partial f^2}{\partial c_t \partial e_{t-1}} = \delta^{t-1} \frac{\partial u^2}{\partial c_t \partial e_{t-1}} \geq 0$ as last period's effort should have no negative effect on this period's consumption. These assumptions, while not trivial, ensure the utility function is reasonably well behaved – otherwise if last period's effort greatly reduced the demand for consumption, it could theoretically reduce this period's effort as well as the demand for consumption has decreased so dramatically.[117] Thus, in conclusion, the conditions are satisfied, and f is supermodular in x.

MS Theorem 4: Using these results from Theorem 6, apply Theorem 4 of Milgrom and Shannon (1995) to achieve the main result. The theorem states that if $f : X \times Z \to \mathbb{R}$, where X is a lattice, T is a partially ordered set, and $S \subset X$, then $argmax_{x \in S} f(x,z)$ is a monotone nondecreasing function in (z, S) if and only if f is quasisupermodular in x and satisfies the single crossing property in $(x; z)$.

MS Theorem 4 Conditions: First note that $\mathbb{R}^n$ with component-wise order forms a lattice as for $\forall x,\, y \in \mathbb{R}^n$, $x \wedge y$ and $x \vee y$ are both in $\mathbb{R}^n$. By the same token, $\mathbb{R}^m$ with component-wise ordering is a partially ordered set. Thus, using the lagrangian function above as f, with x and z defined as above, we have already established supermodularity in x, which implies quasisupermodularity in x. In addition, as f has increasing

[116]Though this assumption of only last period's effort influencing this period is not necessary for this proof. For example, if all previous periods' efforts entered the utitlity function as a discounted sum with non-negative weights, as long as effort is positive with respect to that sum (positive effort momentum), the result would be the same.

[117]For example, if after working a long day, the agent no longer cared for consumption. Under this example, the agent might call in sick, even though effort would have been easier due to effort momentum. In this odd model of behavior, an increase in piece rate last period could decrease effort in the next period.

differences, it satisfies the single crossing property. Thus it follows that $\vec{c}^{*}$, $\vec{e}^{*} \in argmax_{c,e \in \mathbb{R}_{+}^{2T}} f(c,e,z)$ are monotone decreasing functions over $(z,\, \mathbb{R}_{+}^{n})$ – but recall that z_{T+1} was defined as negative γ_1 , z_{T+2} as negative γ_2, and so on. Thus, consumption and effort are monotonically non-decreasing over piece rate vector $\vec{w}$ and monotonically non-increasing over leisure technology vector $\vec{\gamma}$.[118]

Alternate Proofs

It's possible to relax the assumptions of the model. The differentiability of u is unnecessary as long as the conditions of increasing differences / single crossing condition and quasi-supermodularity are satisfied. However, I felt the assumptions above are more familiar with readers compared to assumptions of increasing differences. In addition, the assumption that only last period enters the utility function is unnecessary.

It may also be worth mentioning that there are other ways to achieve a similar result. An earlier draft included a proof using the Multivariate Implicit Function Theorem and also assumed second order conditions and positive determinant Jacobian matrices to get the stronger result of effort strictly increasing in piece rates (or strictly decreasing in leisure technology). However, the matrix notation was cumbersome relative to the above proof.

Proposition 2: Naive Momentum g(.) function

Under the FOC for e_t, c_t:

$$\begin{aligned} -u_e(c_t^*, e_t^*, e_{t-1}, \gamma_t) &= \lambda w_t \\ u_c(c_t^*, e_t^*, e_{t-1}, \gamma_t) &= \lambda p_t \end{aligned}$$

As u is strictly concave over the first argument, this allows for inverse of u_c:

$$c_t^* \quad = \quad u_c^{-1}\left(\lambda p_t,\, e_t^*,\, e_{t-1},\, \gamma_t\right)$$

Which can be inserted into the first FOC to give:

$$-u_e(u_c^{-1}\left(\lambda p_t,\, e_t^*,\, e_{t-1},\, \gamma_t\right), e_t^*,\, e_{t-1}, \gamma_t) \quad = \quad \lambda w_t$$

[118] To be clear, as I am using component-wise ordering, increasing just one element of the piece rate vector or one element of the leisure technology vector causes the vector to be ordered as higher than the unaltered vector.

The e_t^* which solves this first order condition is equivalent to the e_t^{**} which would maximize (by construction):

$$e_t^{**} = argmax_{e_t} \ \lambda w_t e_t + \int_0^{e_t} u_e(u_c^{-1}(\lambda p_t, x, e_{t-1}, \gamma_t), x, e_{t-1}, \gamma_t)dx$$

This objective function can be rewritten as $U_t^{'} = \lambda w_t e_t - g(e_t, e_{t-1}, \gamma_t)$. It remains to be shown that this $g(e_t, e_{t-1}, \gamma_t)$ is convex in e_t, which in this case is equivalent to having a negative second derivative:

$$\begin{aligned} \frac{dg}{de_t} &= -\frac{\partial u(u_c^{-1}(\lambda p_t, e_t^*, e_{t-1}, \gamma_t), e_t^*, e_{t-1}, \gamma_t)}{\partial e_t} \\ \frac{d^2 g}{de_t^2} &= -\frac{\partial^2 u}{\partial e_t^2} - \frac{\partial^2 u}{\partial e_t \partial c_t}\frac{dc_t}{de_t} \end{aligned}$$

Note $\frac{de_{t-1}}{de_t} = \frac{d\gamma_t}{de_t} = 0$ as e_{t-1} and γ_t are not choice variables at time t. The total differential of the first order condition for c_t yields:

$$\begin{aligned} u_{cc}dc_t + u_{ce}de_t &= 0 \\ \Rightarrow \frac{dc_t}{de_t} &= -\frac{u_{ce}}{u_{cc}} \end{aligned}$$

Thus for the second derivative:

$$\begin{aligned} \frac{d^2 g}{de_t^2} &= -u_{ee} + u_{ec}\frac{u_{ce}}{u_{cc}} \\ &= -\frac{1}{u_{cc}}(u_{ee}u_{cc} - u_{ec}u_{ce}) \\ &> 0 \end{aligned}$$

As $u_{cc} < 0$ and $u_{ee}u_{cc} - u_{ec}u_{ce} > 0$. Given this derivation, one can derive how past, present, and future piece rate and leisure technology influence effort as outlined in Section 2.

Proposition 3: Reciprocity Proof 1

If $\alpha_2 > 0$ and e_1, e_2 is an interior solution, then $\frac{\partial e_{t+1}}{\partial \gamma_t} > 0$

$$U = (\lambda w_1 + \alpha_1 w_1 + \alpha_1 w_2 + \alpha_2 \gamma_1 + \alpha_2 \gamma_2)(e_1 + e_2) - g(\gamma_1 e_1) - g(\gamma_2 e_2)$$

First order condition:

$$\begin{aligned} (\lambda w_1 + \alpha_1 w_1 + \alpha_1 w_2 + \alpha_2 \gamma_1 + \alpha_2 \gamma_2) &= \gamma_1 g'(\gamma_1 e_1) \\ (\lambda w_2 + \alpha_1 w_1 + \alpha_1 w_2 + \alpha_2 \gamma_1 + \alpha_2 \gamma_2) &= \gamma_2 g'(\gamma_2 e_2) \end{aligned}$$

Multivariate implicit function theorem (Dini) gives us:

$$
\begin{aligned}
\frac{\partial e_2}{\partial \gamma_1} &= -\frac{det\begin{bmatrix} \frac{\partial F_1}{\partial e_1} & \frac{\partial F_1}{\partial \gamma_1} \\ \frac{\partial F_2}{\partial e_1} & \frac{\partial F_2}{\partial \gamma_1} \end{bmatrix}}{det\begin{bmatrix} \frac{\partial F_1}{\partial e_1} & \frac{\partial F_1}{\partial e_2} \\ \frac{\partial F_2}{\partial e_1} & \frac{\partial F_2}{\partial e_2} \end{bmatrix}} \\
&= -\frac{det\begin{bmatrix} -\gamma_1^2 g^{''}(\gamma_1 e_1) & \alpha_2 - g^{'}(\gamma_1 e_1) - \gamma_1 e_1 g^{''}(\gamma_1 e_1) \\ 0 & \alpha_2 \end{bmatrix}}{det\begin{bmatrix} -\gamma_1^2 g^{''}(\gamma_1 e_1) & 0 \\ 0 & -\gamma_2^2 g^{''}(\gamma_2 e_2) \end{bmatrix}} \\
&= -\frac{-\alpha_2 \gamma_1^2 g^{''}(\gamma_1 e_1)}{\gamma_1^2 \gamma_2^2 g''(\gamma_1 e_1) g''(\gamma_2 e_2)} \\
&= \frac{\alpha_2}{\gamma_2^2 g''(\gamma_2 e_2)} > 0
\end{aligned}
$$

Proposition 3: Reciprocity Proof 2

If $\alpha_1 > 0$ and e_1, e_2 is an interior solution, then $\frac{\partial e_{t+1}}{\partial w_t} > 0$:

By Multivariate Implicit Function Theorem using the above FOC.

$$
\begin{aligned}
\frac{\partial e_2}{\partial w_1} &= -\frac{det\begin{bmatrix} \frac{\partial F_1}{\partial e_1} & \frac{\partial F_1}{\partial w_1} \\ \frac{\partial F_2}{\partial e_1} & \frac{\partial F_2}{\partial w_1} \end{bmatrix}}{det\begin{bmatrix} \frac{\partial F_1}{\partial e_1} & \frac{\partial F_1}{\partial e_2} \\ \frac{\partial F_2}{\partial e_1} & \frac{\partial F_2}{\partial e_2} \end{bmatrix}} \\
&= -\frac{det\begin{bmatrix} -\gamma_1^2 g^{''}(\gamma_1 e_1) & \lambda + \alpha_1 \\ 0 & \alpha_1 \end{bmatrix}}{det\begin{bmatrix} -\gamma_1^2 g^{''}(\gamma_1 e_1) & 0 \\ 0 & -\gamma_2^2 g^{''}(\gamma_2 e_2) \end{bmatrix}} \\
&= -\frac{-\alpha_1 \gamma_1^2 g^{''}(\gamma_1 e_1)}{\gamma_1^2 \gamma_2^2 g''(\gamma_1 e_1) g''(\gamma_2 e_2)} \\
&= \frac{\alpha_1}{\gamma_2^2 g''(\gamma_2 e_2)} > 0
\end{aligned}
$$

A.1.5. Chapter 1 Secondary Outcome Treatment Effects

In addition to correct problems solved as a metric for effort, I also collected the number of YouTube videos searched as a proxy for leisure time.[119] As predicted, subjects in the first experiment search 0.2 fewer YouTube videos when the piece rate is increased (see Appendix Table 5). In some specifications, this reduction in leisure persists into the next period as well. This is consistent with the model in which YouTube videos are a leisurely activity, and when faced with a higher piece rate, the agent exerts more time and effort working (and less on leisure).

To further validate that the phone access was actually a leisure activity, subjects in the first experiment searched about 0.12 to 0.15 fewer searches ($p < 0.05$) when given access to cell phones. This is consistent with a model in which phone access and YouTube videos are both leisure activities that compete for attention. Anecdotally, both YouTube and the cellphone often rely on visual cues on different screens, making them difficult to serve as leisure complements.

A.1.6. Chapter 1 Median Unbiased Estimator (MUE)

In this appendix section, I employ a Median Unbiased Estimator as an alternate specification (as opposed to instrumental variables).

Andrews (1993) outlines a method for adjusting the well known bias of using OLS to estimate an AR(1) for three cases: (i) without an intercept, (ii) with an intercept, and (iii) with an intercept and time trend. Unfortunately, the original Andrews (1993) paper does not allow for individual fixed effects or individual exogenous variables x_{it}. Estimating the model without fixed effects or individual-level covariates would likely bias the ρ parameter upwards through omitted variable bias, as individuals have time-constant heterogeneity in their effort allocation.[120] Thus the Andrews (1993) MUE estimator would still be biased in this setting as it would be too low, unable to account for the additional omitted variable bias.

However, there has been considerable work extending the original MUE estimator to panel data. One direction can be found in work on Panel Exactly Median-Unbiased Estimators (PEMU) by Cermeno (1999) and Phillips and Sul (2003). However, an important assumption of this early work is that the error terms are

[119] Ideally I would prefer total length of YouTube videos played, but this information could be gathered from the web browser. In addition, playing a video does not necessarily imply that the agent is actually watching a video, and length of the video may be imperfect measures if the agents skip portions. Thus, both would be noisy estimates of the total leisure time.

[120] For example, if there were two types of workers, lazy and hard working, then even if the true ρ was 0 in the model above, because the individual heterogeneity is not being addressed, you would detect a very large $\hat{\rho}$.

homoskedastic and i.i.d. normal.[121] In addition, these works do not allow for other exogenous regressors x_{it} aside from individual and time fixed effects. Under these assumptions, the mapping between $\hat{\rho}^{LSDV}$ from Least Squares Dummy Variable (LSDV) and the median unbiased $\hat{\rho}^{MU}$ does not depend on the individual fixed effects and can be obtained by Monte Carlo simulations.

Carree (2002) extends Andrews (1993) by allowing exogenous variables x_{it} to be included as well as individual fixed effects. This addition may be important as I have exogenous treatment variables (piece rates and leisure options) that could influence effort. As Carree (2002) proves, the Least Squares Dummy Variable estimator of ρ will still be biased downward, as in the original Nickell (1981) paper. One additional benefit of the Carree (2002) paper is that it provides closed form solutions for $T = 2$ and $T = 3$, which one of my experiments satisfies.[122] This enabled me to provide you some results below, but does not provide closed forms for the standard errors (which would be estimated by Monte Carlo simulations).

To reiterate, I will be applying the Carree (2002) results using the piece rate as the exogenous variable and number of problems solved (per 5 minutes) as the outcome variable, and differencing out the running sum to remove individual fixed effects:

$$\begin{aligned}
e_{it} &= \rho \cdot e_{it-1} + \mu_i + \beta x_{it} + \epsilon_{it} \\
\Rightarrow e_{it} - \bar{e}_{it} &= \rho \cdot (e_{it-1} - \bar{e}_{i,t-1}) + \beta \cdot (x_{it} - \bar{x}_{it}) + (\epsilon_{it} - \bar{\epsilon}_{it}) \\
\tilde{e}_{it} &= \rho \cdot \tilde{e}_{it-1} + \beta \cdot \tilde{x}_{it-1} + \tilde{\epsilon}_{it}
\end{aligned}$$

As in Nickell (1981) and proven in Carree (2002), the ρ estimated from the OLS of this specification is still biased downward. However, Carree provides a median unbiased estimator when $T = 3$ (as in my case). Specifically:

---

[121] These assumptions are relaxed in Phillips and Sul (2003) in an estimator called Panel Feasible Generalized Least Squares Median-Unbiased Estimators (PFGLSMUE). A less important difference is that they assume the AR(1) component is actually embedded in a latent variable:

$$\begin{aligned}
e_{it} &= \mu_i + e^*_{it} + \epsilon_{it} \\
e^*_{it} &= \rho e^*_{it-1} + \nu_{it}
\end{aligned}$$

However, this can still be mapped to my original model above by scaling $\mu^*_i = \mu_i/(1-\rho)$

[122] Although the paper only gets "nearly" unbiased asymptotic estimators for $T > 3$, for $T = 2$ and $T = 3$ the estimator is asymptotically exactly unbiased. However, the usual issues with sample size remains.

$$\hat{\rho}^{MUE} = \frac{9\hat{\rho}^{OLS} + 2\hat{g}}{9 - \hat{g}} \qquad \text{(from equation 12b)}$$

$$\hat{g} \equiv \frac{\hat{\sigma}^2_{\tilde{\epsilon}}}{(1 - \text{corr}_{\tilde{x},\tilde{y}_{t-1}}) \cdot \hat{\sigma}^2_{\tilde{y}_{t-1}}} \qquad \text{(from equation 10)}$$

When I constructed the above OLS regression differincing out running means, I received a $\hat{\rho}^{OLS} = 0.1052$. Once I use this method for correcting the bias, I receive a $\hat{\rho}^{MUE} = 0.4610$ which is very close to the Instrumental Variable estimates I find in the paper of 0.43 to 0.45. I provide more details below on how this estimate was constructed:

Term	Estimate	Origin
$\hat{\sigma}^2_{\tilde{\epsilon}}$	2.686	Estimated from residuals of OLS of differenced equation above
$\hat{\sigma}^2_{\tilde{y}_{t-1}}$	2.066	Estimated from the data of $\tilde{y}_{i,t-1}$
$\text{corr}_{\tilde{x},\tilde{y}_{t-1}}$	5.23×10^{-7}	Estimated from the data of $\tilde{x}_{i,t}$, $\tilde{y}_{i,t-1}$. Note this should be 0 (see details below).
$\hat{g}$	1.301	Transformation of above statistics using closed form solution
$\hat{\rho}^{OLS}$	0.1052	Estimated coefficient from OLS of differenced equation above (biased downward)
$\hat{\rho}^{MUE}$	0.4610	$= (9\hat{\rho}^{OLS} + 2\hat{g})/(9 - \hat{g})$ from Carree (2002), $T = 3$ case, equation 12b
$\hat{\rho}^{IV}$	0.43 to 0.45	From IV strategy, see Table 5

Note that $\text{corr}_{\tilde{x},\tilde{y}_{t-1}}$ is essentially 0. This is not a mistake or a sign of a weak regressor, but rather a sign that treatment was properly randomized. Because it represents the correlation between $\tilde{x}$ at time t and $\tilde{y}$ at time t-1, this is saying that last period's effort difference does not predict the piece rate treatment in the next period. This is to be expected as treatment was randomized, so last period's effort should not predict next period's piece rate treatment.

As Carree (2002) mentions, "An exogenous variable which is very highly correlated with the lagged endogenous variable and which provides little additional explanatory power will lead to worse bias." An exogenous variable x which is highly correlated with the lagged endogenous variable would result in a large $\text{corr}_{\tilde{x},\tilde{y}_{t-1}}$. As can be seen from the above equations, a large correlation would *increase* $\hat{g}$, increasing the bias. However a predictive x also helps lower $\sigma^2_{\tilde{\epsilon}}$, which reduces the bias. Thus, there is a potential trade off for including exogenous covariates. In my case, piece rate helps predict effort – the $\hat{\beta}$ from the above regression was 11.99 with a t-value of 5.49, very much in line with the Instrumental Variable results from the main specification – and does not correlate with past effort, so it helps reduce bias in both directions.

However, there are two potential issues to address with this estimation. First, as with PEMU models, this model's bias correction relies on homoskedastic and i.i.d. normal errors. Second, as with other MUE bias correction, the estimator is only median unbiased asymptotically, as $N \to \infty$. While Monte Carlo results have explored the small sample properties of these estimators in some cases, this is still a potential concern.

A.1.7. Chapter 2 Appendix Tables

Appendix Table 1. Credit and \$100: Impact on Mean of Selection Value

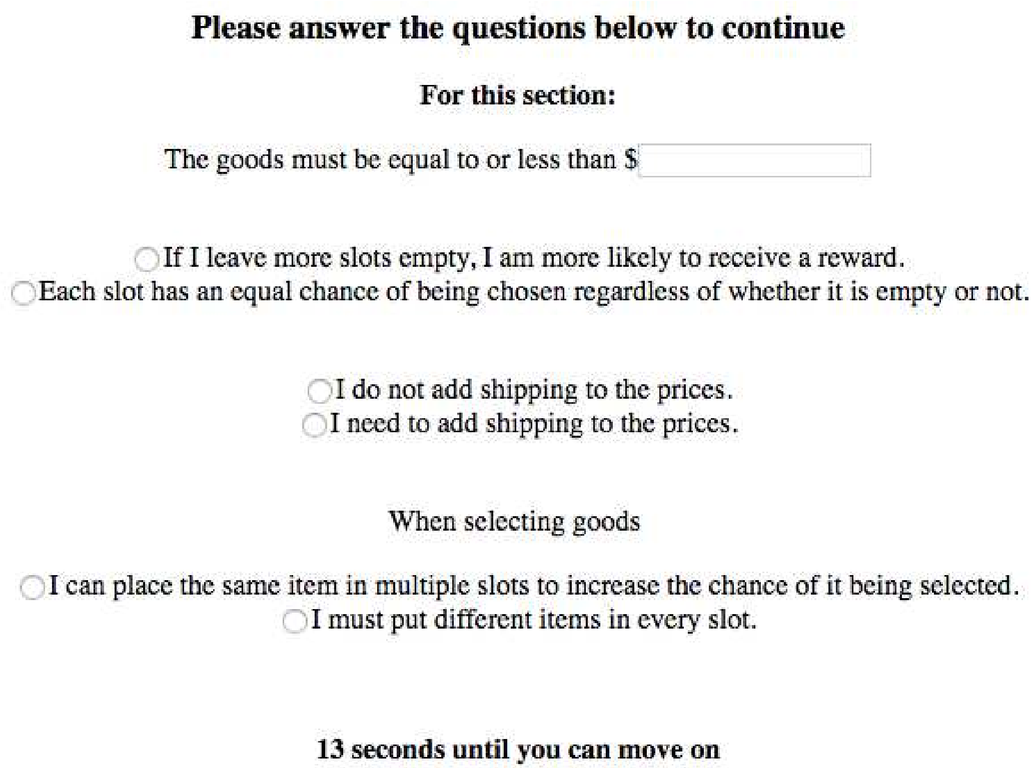
$$AverageValue_{i,t} \quad = \quad \alpha \cdot Credit_{i,t} + \beta \cdot 100Treatment_{i,t} + \gamma X_i + \epsilon_{i,t}$$

Dependent Variable	Specification			
Average Value	(1)	(2)	(3)	(4)
Subject Selects Credit	0.56***	0.57***	0.52***	0.52***
(Binary Treatment Var.)	(0.12)	(0.12)	(0.12)	(0.13)
\$100 Total Allocation	0.29***	0.30***	0.30***	0.30***
(Binary Treatment Var.)	(0.11)	(0.11)	(0.11)	(0.11)
First Period		0.12	0.12	0.12
		(0.11)	(0.11)	(0.11)
Session Fixed Effects			X	X
Individual Controls				X
Dependent Variable Mean	9.32	9.32	9.32	9.32
Number of Observations	248	248	248	248
Number of Individuals	124	124	124	124
Adj-R^2	0.09	0.10	0.12	0.17

Notes. The dependent variable is the average value of the entries in a single period. All specifications report results from OLS regressions and also include a constant term. Individual Controls include sex, age, ethnicity bins, and number of sessions done. Standard errors are given in parentheses and clustered at the subject (individual) level. $* = p < 0.1$, $** = p < 0.05$, $*** = p < 0.01$.

A.1.8. Chapter 2 Appendix Figures

Appendix Figure 1. Quiz for Introduction Instructions



Please answer the questions below to continue

For this section:

The goods must be equal to or less than $

If I leave more slots empty, I am more likely to receive a reward.
Each slot has an equal chance of being chosen regardless of whether it is empty or not.

I do not add shipping to the prices.
I need to add shipping to the prices.

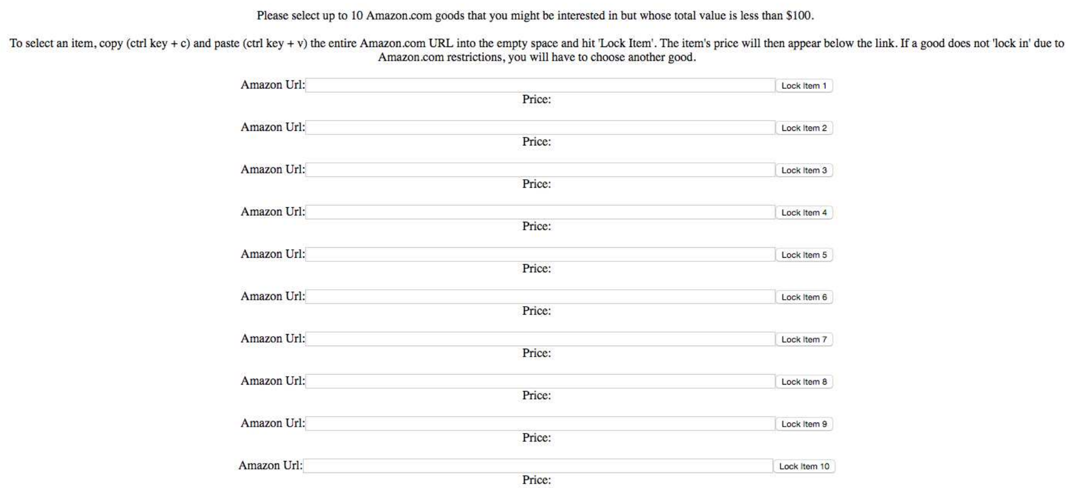
When selecting goods

I can place the same item in multiple slots to increase the chance of it being selected.
I must put different items in every slot.

13 seconds until you can move on

Notes. Every participant had to answer questions after reading experiment instructions. Subjects had to answer all questions correctly to proceed. If the subject entered the wrong answers, the browser would alert them to this and ask for them to review the instructions again.

Appendix Figure 2. Quiz for Instructions Prior to Each Period


Please select up to 10 Amazon.com goods that you might be interested in but whose total value is less than $100.

To select an item, copy (ctrl key + c) and paste (ctrl key + v) the entire Amazon.com URL into the empty space and hit 'Lock Item'. The item's price will then appear below the link. If a good does not 'lock in' due to Amazon.com restrictions, you will have to choose another good.

Amazon Url: Lock Item 1
Price:

Amazon Url: Lock Item 2
Price:

Amazon Url: Lock Item 3
Price:

Amazon Url: Lock Item 4
Price:

Amazon Url: Lock Item 5
Price:

Amazon Url: Lock Item 6
Price:

Amazon Url: Lock Item 7
Price:

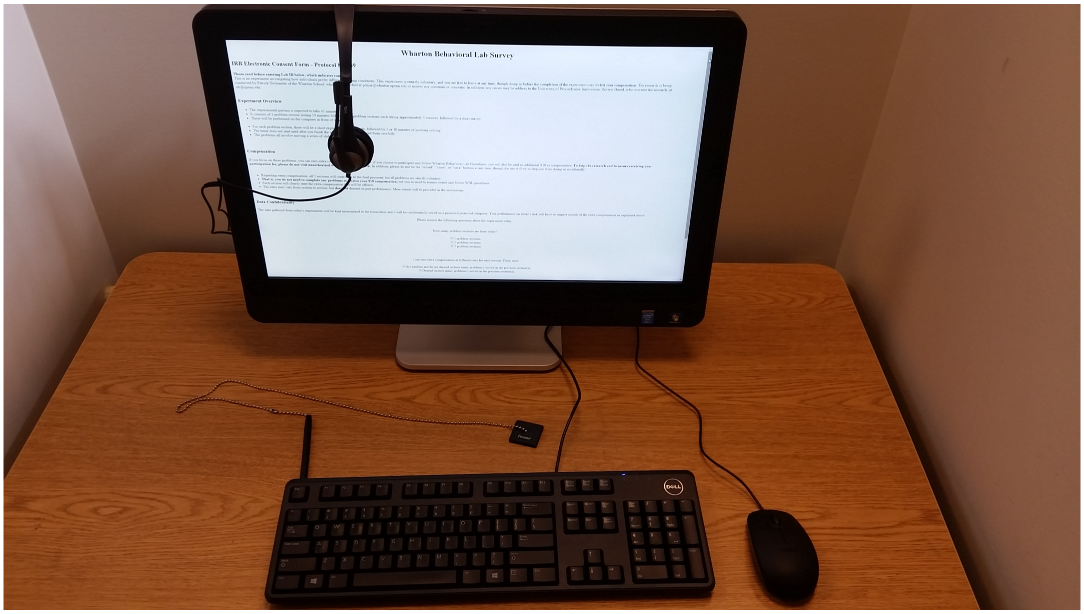
Amazon Url: Lock Item 8
Price:

Amazon Url: Lock Item 9
Price:

Amazon Url: Lock Item 10
Price:

Notes. Every participant had to answer questions prior to every period. If the subject entered the wrong answers, the browser would alert them to this and ask for them to review the instructions again.

Appendix Figure 3. Cubicle Environment



Notes. Every participant had access to an identical computer with headphones as pictured above. Cookies and browser history were cleared after every session to limit any subject overlap. It was not possible to see other subjects from within the cubicle. Google Chrome was employed as the browser. All instructions were written, but lab assistants were on site to answer any additional questions.

A.1.9. Chapter 3 Proofs

Proof of Proposition 3.2.2. The sequence $\{D(t)\}$ is monotone and bounded. Thus, by the Monotone Convergence Theorem, it converges to some number, say $\overline{d} \geq 0$. We need to show that the sequence $\{D(t+1) + D(t-1) - 2D(t)\}$ has no negative limit points:

$$\liminf_{t\to\infty} \left(D(t+1) + D(t-1) - 2D(t)\right) \geq 0.$$

Suppose this is not true. Then there exists $\epsilon > 0$ and a subsequence $\{D(t_k)\}$ such that

$$D(t_k+1) + D(t_k-1) - 2D(t_k) \leq -\epsilon$$

for all t_k. However, because $D(t_k)$ converges to $\overline{d}$, it follows that $D(t_k+1) + D(t_k-1) - 2D(t_k)$ converges to zero. Thus, there exists $\bar{t}_k$ such that for all $t > \bar{t}_k$,

$$-\frac{\epsilon}{2} \leq D(t_k+1) + D(t_k-1) - 2D(t_k) \leq \frac{\epsilon}{2},$$

which contradicts the previous inequality. ■

Proof of Proposition 3. The proof will use a couple of Lemmas.

First notice that, because preferences are dynamically consistent, there is no loss in taking $t = 3$. To simplify the expressions, it is convenient to write $\lambda \equiv (c+x)/c > 1$ to denote the consumption with the prize as a proportion of consumption without it. Using the formula in the text, the utility of the safe lottery equals

$$V_0 = \left[(1-\beta)\, c\right]^{\frac{1}{1-\rho}} \cdot \left[1 + \beta + \lambda^{1-\rho}\beta^2 + \frac{\beta^3}{1-\beta}\right]^{\frac{1}{1-\rho}},$$

and the utility of the risky lottery is

$$V_0 = \left[(1-\beta)\, c\right]^{\frac{1}{1-\rho}} \left\{ 1 + \beta \left[\frac{\left(\lambda^{1-\rho} + \frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}} + \left(1 + \beta + \lambda^{1-\rho}\beta^2 + \frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2} \right]^{\frac{1-\rho}{1-\alpha}} \right\}^{\frac{1}{1-\rho}}.$$

Therefore, preferences are locally RSTL at t if and only if the following inequality holds:

$$\left\{1+\beta\left[\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}\right]^{\frac{1-\rho}{1-\alpha}}\right\}^{\frac{1}{1-\rho}} > \left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1}{1-\rho}}. \tag{A.1}$$

To simplify notation, let $f(x) \equiv x^{\frac{1-\alpha}{1-\rho}}$. In the proofs, we will repeatedly use the following result. The expected discounted payoff from the risky lottery exceeds the one from the safe lottery if and only if the intertemporal elasticity of substitution exceeds 1. Formally:

$$\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2}\left\{\begin{array}{c} > \\ < \end{array}\right\}1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta} \iff \rho\left\{\begin{array}{c} < \\ > \end{array}\right\}1. \tag{A.2}$$

We first verify that (A.1) always holds when $\alpha \leq 1$.

Lemma A.1.1 *Let $\alpha \leq 1$. Then, preferences are RSTL.*

Proof There are three cases: (i) $\alpha \leq \rho \leq 1$, (ii) $\rho < \alpha \leq 1$, and (iii) $\alpha \leq 1 < \rho$.

Case i: $\alpha \leq \rho \leq 1$. Since $1-\rho < 0$, inequality (A.1) can be written as

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2} > \left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}.$$

Algebraic manipulations establish that the expected discounted payment of the risky lottery exceeds the one from the safe lottery. Because $\rho < 1$, inequality (A.2) gives

$$\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2} > 1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}.$$

The result then follows from Jensen's inequality since $f(x)$ is increasing and convex when $\alpha, \rho \leq 1$.

Case ii: $\rho < \alpha \leq 1$. To simplify notation, perform the following change of variables: $\gamma \equiv \frac{1-\alpha}{1-\rho} \in (0,1)$ where $\gamma > 0$ since both α and ρ are lower than 1, and $\gamma < 1$ since $\alpha > \rho$. We can rewrite inequality (A.1) substituting α for γ as

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\gamma}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\gamma}}{2} > \left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\gamma}.$$

Rearrange this condition as

$$\left(\frac{1}{\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta}\right)^{\gamma}+\left(\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta\right)^{\gamma}>2.$$

It is straightforward to verify that the expression on the left ("LHS") is a convex function of γ. Recall that $\gamma \in (0,1)$. Evaluating at $\gamma = 0$, we obtain

$$LHS|_{\gamma=0}=2.$$

Since LHS is a convex function of γ, it suffices to show that its derivative wrt γ at zero is positive. We claim that this is true. To see this, notice that

$$\left.\frac{dLHS}{d\gamma}\right|_{\gamma=0}=\ln\left(\frac{\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta}{\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta}\right), \tag{A.3}$$

which, with some algebraic manipulations, can be shown to be strictly positive for any $\rho < 1$. Thus, $LHS > 2$ for all $\gamma \in (0,1]$, establishing RSTL.

Case iii: $\alpha \leq 1 < \rho$. Inequality (A.1) can be simplified as

$$\left[\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}+\frac{\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}\right]^{\frac{1-\rho}{1-\alpha}}<1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}.$$

Since $\frac{1-\alpha}{1-\rho}<0$, this holds if

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}>\left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}. \tag{A.4}$$

Notice that $f(x)=x^{\frac{1-\alpha}{1-\rho}}$ is convex since

$$f''(x)=\left(\frac{1-\alpha}{1-\rho}\right)\left(\frac{1-\alpha}{1-\rho}-1\right)x^{\frac{1-\alpha}{1-\rho}-2}>0,$$

where we used $\frac{1-\alpha}{1-\rho}<0$ and $\frac{1-\alpha}{1-\rho}-1<0$. Thus, by Jensen's inequality,

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}>\left(\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2}\right)^{\frac{1-\alpha}{1-\rho}}. \tag{A.5}$$

From condition (A.2), we have

$$\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2}<1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}.$$

Raising to $\frac{1-\alpha}{1-\rho}<0$, gives

$$\left(\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2}\right)^{\frac{1-\alpha}{1-\rho}}>\left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}.$$

Substituting in (A.5), we obtain

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}>\left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}},$$

which is precisely the condition for RSTL (A.4).

Lemma A.1.2 *Let $\alpha\leq\rho$. Then, preferences are RSTL.*

Proof By Lemma A.1.1, the result is immediate when $\alpha\leq 1$. Therefore, let $\alpha>1$ (which, by the statement of the lemma, requires $\rho>1$).

Rearranging inequality (A.1), we obtain the following condition for RSTL:

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}+\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}}{2}<\left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\frac{1-\alpha}{1-\rho}}. \tag{A.6}$$

Moreover, from condition (A.2), we have

$$\frac{\lambda^{1-\rho}+\frac{\beta}{1-\beta}+1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}}{2}<1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}.$$

Notice that $f(x)$ is increasing when $\alpha,\rho\geq 1$ and it is concave when $\rho\geq\alpha$. Then, condition (A.6) follows by Jensen's inequality.

We are now ready to prove the main result:

Proof of Proposition 3.3.1 First, suppose $\rho<1$. Let $\gamma\equiv-\frac{1-\alpha}{1-\rho}\in(0,+\infty)$ so we can rewrite inequality (A.1) in terms of γ and ρ as

$$\frac{1}{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\gamma}}+\frac{1}{\left(1+\beta+\lambda^{1-\rho}\beta^2+\frac{\beta^3}{1-\beta}\right)^{\gamma}}<\frac{2}{\left(1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}\right)^{\gamma}},$$

which can be simplified as:

$$\left(\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta\right)^{\gamma}+\left(\frac{1}{\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta}\right)^{\gamma}<2.$$

The first term in the expression on the left ("LHS") is convex and decreasing in γ, because the term inside the first brackets is smaller than 1:

$$\rho\leq 1 \implies \frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta\leq 1$$

The second term is convex and increasing in γ because the term inside the second brackets is greater than 1:

$$\rho\leq 1 \implies \frac{1}{\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta}\geq 1.$$

Since the sum of convex functions is convex, it follows that LHS is a convex function of γ.

Evaluating γ at the extremes, we obtain:

$$LHS|_{\gamma=0}=\left(\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta\right)^{0}+\left(\frac{1}{\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta}\right)^{0}=2,$$

and

$$\lim_{\gamma\to\infty} LHS=+\infty>2.$$

Moreover, we claim that the derivative of the LHS wrt γ at zero is negative. To see this, note that

$$\left.\frac{dLHS}{d\gamma}\right|_{\gamma=0}=\ln\left(\frac{\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta}{\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^2}{1-\beta}}+\beta}\right),$$

which, following some algebraic manipulations, can be shown to be strictly negative.

Thus, there exists $\bar{\gamma}>0$ such that $LHS>2$ (RATL) if and only if $\gamma>\bar{\gamma}$. But, since $\gamma\equiv-\frac{1-\alpha}{1-\rho}$ (so that γ is strictly increasing in α), this establishes that there exists a finite $\bar{\alpha}_{\rho,\beta}>\max\{1,\rho\}$ such that we have RATL if $\alpha>\bar{\alpha}_{\rho,\beta}$ and RSTL if $\alpha<\bar{\alpha}_{\rho,\beta}$. This concludes the proof for $\rho<1$.

Now suppose that $\alpha>\rho\geq 1$ (the result is trivial if $\alpha\leq\rho$ from Lemma A.1.2). Let $\gamma\equiv\frac{1-\alpha}{1-\rho}\geq 1$. Then, we have RSTL if and only if

$$\frac{\left(\lambda^{1-\rho}+\frac{\beta}{1-\beta}\right)^{\gamma}+\left(1+\beta+\lambda^{1-\rho}\beta^{2}+\frac{\beta^{3}}{1-\beta}\right)^{\gamma}}{2}<\left(1+\lambda^{1-\rho}\beta+\frac{\beta^{2}}{1-\beta}\right)^{\gamma}.$$

Rearrange this condition as

$$\left(\frac{1}{\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta}\right)^{\gamma}+\left(\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^{2}}{1-\beta}}+\beta\right)^{\gamma}<2. \tag{A.7}$$

As before, it can be shown that the expression on the left ("LHS") is a convex function of γ. Notice that $\lim_{\gamma\to\infty} LHS=+\infty>2$. Morevoer, $LHS|_{\gamma=1}<2$ since, with some algebraic manipulations, one can show that

$$\lambda^{1-\rho}<1 \iff \frac{1}{\frac{1}{\lambda^{1-\rho}+\frac{\beta}{1-\beta}}+\beta}+\frac{1}{1+\lambda^{1-\rho}\beta+\frac{\beta^{2}}{1-\beta}}+\beta<2.$$

Thus, there exists $\bar{\gamma}>0$ such that $LHS>2$ (RATL) if and only if $\gamma>\bar{\gamma}$. The result then follows from the fact that γ is increasing in α.

To conclude the proof, it remains to be shown that $\lim_{x\searrow 0}\bar{\alpha}_{\rho,\beta,x}=+\infty$. Both sides of (A.1) are equal to $\left(\frac{1}{1-\beta}\right)^{\frac{1}{1-\rho}}$ when $\lambda=1$. The derivative of the expression on the right (utility of the safe lottery) with respect to λ at $\lambda=1$ is

$$\left(\frac{1}{1-\beta}\right)^{\frac{\rho}{1-\rho}}\beta^{2}. \tag{A.8}$$

The derivative of the expression on the left (utility of the risky lottery) with respect to λ at $\lambda=1$ is

$$\beta\frac{1+\beta^{2}}{2}\left(\frac{1}{1-\beta}\right)^{\frac{\rho}{1-\rho}}. \tag{A.9}$$

With some algebraic manipulations, it can be shown that for any $\beta\in(0,1)$, the term in (A.8) is lower than the one in (A.9).

Proof of Proposition 3.3.2. The first claim was proved in the text. For the second claim, it is enough to show that there is a specific time lottery that will always be preferred to the safe lottery independently of x. For $k\leq t$ and payment x, consider the time lottery $\langle 0.5,(x,t-k);0.5,(x,t+k)\rangle\in\mathcal{P}_x$. Using the formula of DPWU, we have

$$\begin{aligned}
V_{DPWU}\left(\delta_{(x,t)}\right) &\geq V_{DPWU}\left(\langle 0.5,(x,t-k);0.5,(x,t+k)\rangle\right)\Leftrightarrow\\
D(t) &\geq \pi(0.5)\,D(t-k)+(1-\pi(0.5))\,D(t+k)
\end{aligned}$$

Take $k = t$ and recall that $D(0) = 1$, this holds if and only if

$$D(k) \geq \pi\,(0.5) + (1 - \pi\,(0.5))\, D(2k)$$

or

$$D(k) - D(2k) \geq \pi\,(0.5)\,(1 - D(2k)) \tag{A.10}$$

But D is a decreasing function which is by 0. By the the monotone convergence theorem, $\lim\limits_{k\to\infty} D(k) - D(2k) = 0$, while the right hand side of equation (A.10) is bounded below by

$\pi\,(0.5)\left(1 - \lim\limits_{k\to\infty} D(2k)\right) > 0.$ ■

A.1.10. RATL with Consumption Smoothing

In this appendix, we consider the choice between safe and risky time lotteries when the decision maker can freely save and borrow. While the main benchmark is the Discounted Expected Utility model (DEU), we will consider the more general Epstein-Zin model (EZ) we discuss in Section 3. EZ coincides with DEU when $\alpha = \rho$.

We study a standard consumption-savings model with no liquidity constraints. The decision maker allocates income between consumption and a riskless asset that pays a constant interest rate r. Let $D \equiv \frac{1}{1+r} > 0$ denote the market discount rate. In period t, the decision maker earns an income W_t. Let $\mathcal{W} \equiv \sum_{t=0}^{\infty} D^t W_t$ denote the net present value of lifetime income (in the absence of the time lottery). For existence, we assume that $\beta < D^{1-\rho}$, which always holds if $\rho \geq 1$.

As in the text, the decision maker faces a choice between a time lottery that pays \$$x$ in period t with certainty ("safe lottery") and a lottery that pays \$$x$ at either $t-1$ or $t+1$ with equal probabilities ("risky lottery"). We will determine the qualitative and quantitative ability of this model to reconcile a preference for the safe lottery. Our qualitative result states that the safe lottery is preferred if people are sufficiently risk averse and, moreover, as the prize decreases, the amount of risk aversion needed to make someone prefer the safe lottery goes to infinity. More precisely:

Proposition A.1.3 *There exists a unique $\bar{\alpha}_{x,D,\mathcal{W}} > 1$ such that the safe time lottery is preferred if and only if $\alpha > \bar{\alpha}_{x,D,\mathcal{W}}$. Moreover, $\lim_{x \searrow 0} \bar{\alpha}_{x,D,\mathcal{W}} = +\infty$.*

Proof It is helpful to consider the utility of deterministic streams of payments first. With deterministic incomes, the optimal consumption solves the following program:

$$\begin{array}{rl} \max_{\{C_t\}} & \left[(1-\beta)\sum_{t=0}^{\infty}\beta^t C_t^{1-\rho}\right]^{\frac{1}{1-\rho}} \\ \text{subject to} & \sum D^t C_t = \mathcal{W} \end{array} .$$

A variational argument establishes the following necessary optimality condition:

$$C_t = \mathcal{W}\left(1 - D^{1-\frac{1}{\rho}}\beta^{\frac{1}{\rho}}\right)\left(\frac{\beta}{D}\right)^{\frac{t}{\rho}} . \tag{A.11}$$

Therefore, the utility from a deterministic stream of payments with net present value $\mathcal{W}$ is

$$\begin{aligned} \mathcal{V}(\mathcal{W}) &\equiv \left\{(1-\beta)\sum_{t=0}^{\infty}\beta^t\left(\mathcal{W}\left(1-D^{1-\frac{1}{\rho}}\beta^{\frac{1}{\rho}}\right)\left(\tfrac{\beta}{D}\right)^{\frac{t}{\rho}}\right)^{1-\rho}\right\}^{\frac{1}{1-\rho}} \\ &= \mathcal{W}\left(1-D^{1-\frac{1}{\rho}}\beta^{\frac{1}{\rho}}\right)\left\{(1-\beta)\left[\sum_{t=0}^{\infty}\left(\tfrac{\beta}{D^{1-\rho}}\right)^{\frac{t}{\rho}}\right]\right\}^{\frac{1}{1-\rho}} . \end{aligned}$$

Notice that this expression is finite if and only if $\beta < D^{1-\rho}$, which we assumed to be the case.

Recall that the safe time lottery pays x in period t and the risky lottery that pays x at either $t-1$ or $t+1$. By dynamic consistency, the choice between these lotteries does not depend on t. For notational simplicity, we therefore set $t = 2$. The utility from the safe lottery is

$$\mathcal{V}(\mathcal{W} + Dx) = (\mathcal{W} + D\Delta)\left(1 - D^{1-\frac{1}{\rho}}\beta^{\frac{1}{\rho}}\right)\left[\frac{1-\beta}{1-\left(\frac{\beta}{D^{1-\rho}}\right)^{\frac{1}{\rho}}}\right]^{\frac{1}{1-\rho}}.$$

The utility from the risky lottery is

$$\left\{\frac{[\mathcal{V}(\mathcal{W}+x)]^{1-\alpha} + \left[\mathcal{V}(\mathcal{W}+D^2x)\right]^{1-\alpha}}{2}\right\}^{\frac{1}{1-\alpha}}$$

$$= \left(1 - D^{1-\frac{1}{\rho}}\beta^{\frac{1}{\rho}}\right)\left[\frac{1-\beta}{1-\left(\frac{\beta}{D^{1-\rho}}\right)^{\frac{1}{\rho}}}\right]^{\frac{1}{1-\rho}}\left[\frac{(\mathcal{W}+x)^{1-\alpha} + \left(\mathcal{W}+D^2x\right)^{1-\alpha}}{2}\right]^{\frac{1}{1-\alpha}}.$$

Comparing these two expressions, it follows that the risky lottery is preferred if and only if

$$\left[\frac{(\mathcal{W}+x)^{1-\alpha} + \left(\mathcal{W}+D^2x\right)^{1-\alpha}}{2}\right]^{\frac{1}{1-\alpha}} \geq \mathcal{W} + Dx. \tag{A.12}$$

Notice that this condition relies only on risk aversion, not on the elasticity of intertemporal substitution. That is, disentangling IES from risk aversion clarifies that only risk aversion matters for the choice between the safe and the risky time lottery.

First, we claim that the risky lottery is always preferred when $\alpha < 1$. To see this, rewrite condition (A.12) as

$$\left(\frac{\mathcal{W}+x}{\mathcal{W}+Dx}\right)^{\xi} + \left(\frac{\mathcal{W}+D^2x}{\mathcal{W}+Dx}\right)^{\xi} \geq 2,$$

where $\xi \equiv 1-\alpha \in [0,1]$. The expression on the left is a convex function of ξ. The result then follows from the fact that, at $\xi = 0$, the inequality holds and that the expression on the left is increasing. Evaluating the expression on the left at $\xi = 0$, gives

$$\left(\frac{\mathcal{W}+x}{\mathcal{W}+Dx}\right)^{0} + \left(\frac{\mathcal{W}+D^2x}{\mathcal{W}+Dx}\right)^{0} = 2. \tag{A.13}$$

The derivative of the expression on the left at $\xi = 0$ equals

$$\ln\left[\frac{(W+x)\left(W+D^2x\right)}{(W+Dx)^2}\right] \geq 0,$$

where the inequality follows from standard algebraic manipulations.

Next, let $\alpha > 1$ and rewrite condition (A.12) as

$$\left(\frac{\mathcal{W} + Dx}{\mathcal{W} + x}\right)^{\psi} + \left(\frac{\mathcal{W} + Dx}{\mathcal{W} + D^2 x}\right)^{\psi} \leq 2,$$

where $\psi \equiv \alpha - 1 > 0$. We claim that there exists a unique interior cutoff such that the inequality holds if and only if ψ lies below this cutoff. Notice that the expression on the left is again a convex function ψ. Moreover, it equals 2 at $\psi = 0$ and it converges to $+\infty$ as $\xi \to \infty$. It suffices to show that the derivative at zero is negative. The derivative of the expression on the left $\psi = 0$ is

$$\ln\left(\frac{(\mathcal{W} + Dx)^2}{(\mathcal{W} + x)(\mathcal{W} + D^2 x)}\right),$$

which can be shown to be negative.

To show that $\lim_{x\to 0} \bar{\alpha}_{x,D,\mathcal{W}} = +\infty$, notice that, at $x = 0$, both sides of (A.13) equal 2. Moreover, tedious algebra establishes that the derivative of the the expression on the left of (A.13) with respect to x evaluated at 0 is negative.

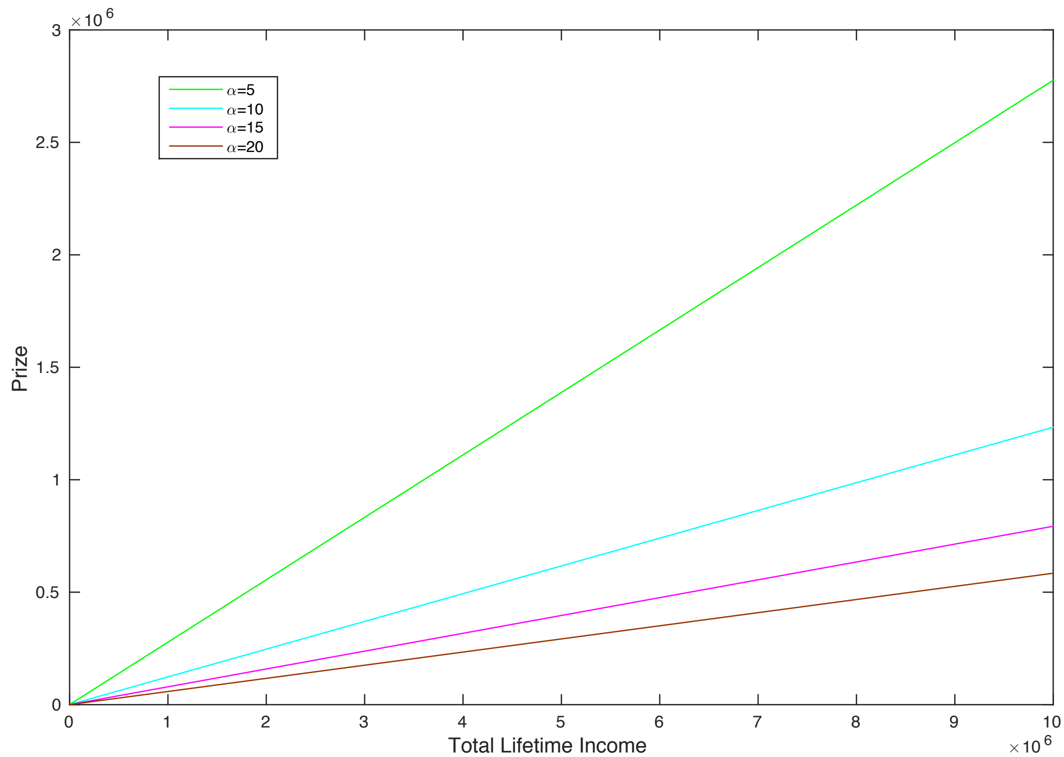
Next, we turn to the quantitative ability of this model to generate a preference for the safe time lottery. Since the condition for the safe lottery to be chosen depends on the risk aversion parameter but not on the elasticity of intertemporal substitution, all results also hold for DEU.

Rationalizing a preference for the safe lottery requires either unreasonably high levels of risk aversion or unreasonably low lifetime incomes. For example, with $D = 0.9$ and $\alpha = 10$ and a net present value of lifetime income of one million dollars, a person would only prefer the safe lottery if the prize exceeded \$123,500!

Figure A.1.10 shows that this is a general pattern. It represents, for each lifetime income (horizontal axis), the prize that would make the individual indifferent between the risky and the safe time lotteries. The risky lottery is preferred if the prize lies below the depicted line, and the safe lottery is preferred if it lies above it. For $\alpha = 5$, the risky lottery is preferred as long as the prize does not exceed 27.7% of the total lifetime income ($\mathcal{W}$). For $\alpha = 10$, the risky lottery is chosen as long as the prize does not exceed 12.3% of $\mathcal{W}$. Even for $\alpha = 25$, a high risk aversion coefficient, the risky lottery is chosen for any prize below 4.6% of $\mathcal{W}$.

Thus, for moderate prizes (including any of the ones in our experiments) and reasonable risk aversion parameters, EZ with smoothing predicts a preference for the risky lottery.

Appendix Figure 1. Indifference regions for risk



Notes. Regions of indifference between the safe and the risky time lotteries, with total lifetime incomes $\mathcal{W}$ on the horizontal axis and prizes x on the vertical axis, for different coefficients of relative risk aversion α (and discount parameter $D = .9$). The risky lottery is preferred at points below the line, and the safe lottery is preferred at points above it.

A.1.11. Additional Experimental Analysis

Table 31: Proportion of RATL subjects

Sample	No Cert. Bias (12-13)		No Cert. Bias (12-14)	
Treatment	Long	Short	Long	Short
Question 1	67.50	55.00	62.50	54.17
Majority in Q1-5	68.75*	50.00	63.89	47.22
MPL in Q10	47.50	51.90	45.83	50.00
MPL in Q11	54.43	48.10	56.94	48.61
Observations	80	80	72	72

Notes. Same as Table 28, including certainty bias measure from Questions 12 and 14 (see footnote 111).

Table 32: Probit Regressions: RATL and Atemporal Risk Aversion

Dep. Var.	RATL Q.1				RATL Majority Q.1-5			
Treatment	Long		Short		Long		Short	
(Probit)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cert. Bias		-.19		.18		-.21		.12
		(-1.56)		(1.20)		(-1.71)		(0.82)
Convexity	3.73*		-4.17		-.39		-10.60***	
	(1.68)		(-1.25)		(-0.19)		(-2.83)	
Constant	.22	.40***	.25*	.19	.41**	.37***	.15	.01
	(1.35)	(2.93)	(1.67)	(1.41)	(2.50)	(2.75)	(1.02)	(.09)
Pseudo-R^2	.02	.02	.01	.01	.01	.03	.07	.01
Obs.	101	95	88	88	101	95	88	88

Notes. Same as Table 29. Each regression excludes one dependent variable.

Table 33: Probit Regressions: RATL and Convexity and Certainty Bias

Dep. Var.	RATL Q.10				RATL Q.11			
Treatment	Long		Short		Long		Short	
(Probit)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Convexity	3.63*		-1.59		1.13		-7.59**	
	(1.73)		(-2.83)		(-0.53)		(-2.20)	
Cert. Bias		-.06		.45***		.19*	.18	
		(-0.52)		(2.84)		(1.71)		(1.23)
Constant	-.29*	-.11	.13	.13	.12	.24*	.12	.18
	(-1.79)	(-.88)	(.88)	(.92)	(0.75)	(1.80)	(.81)	(1.23)
Pseudo-R^2	.02	.01	.01	.07	.01	.02	.04	.01
Obs.	101	95	87	87	101	95	87	87

Notes. Same as Table 29. Each regression excludes one dependent variable.

A.1.12. Chapter 3 Question Example

The following is an example of the questionnaire used in the experiment, in the Short treatment, followed by the instructions used in the experiment. For a full set of questionnaires, please contact the author.

A.1.13. Questionnaire Part 1

QUESTIONNAIRE – PART I

Please indicate your lab id: _________

Please answer each of the following questions by checking the box of the preferred option.

If the question is selected for payment, you will get the payment specified above the question, with a payment date based on your choice and, in some cases, on chance.

Question 1

Payment: **$20.** Payment date:

Option A			**Option B**
2 weeks	☐	☐	75% chance of 1 week 25% chance of 5 weeks

Question 2

Payment: **$15.** Payment date:

Option A			**Option B**
3 weeks	☐	☐	50% chance of 1 week 50% chance of 5 weeks

BIBLIOGRAPHY

J. Abeler, A. Falk, L. Goette, and D. Huffman. Reference points and effort provision. *The American Economic Review*, 101(2):pp. 470–492, 2011. ISSN 00028282. URL http://www.jstor.org/stable/29783680.

M. Agranov and P. Ortoleva. Stochastic choice and preferences for randomization. mimeo California Institute of Technology, 2015.

G. A. Akerlof. The Economics of Caste and of the Rat Race and Other Woeful Tales. *The Quarterly Journal of Economics*, 90(4):599–617, November 1976. URL http://ideas.repec.org/a/tpr/qjecon/v90y1976i4p599-617.html.

M. Allais and G. Hagen. *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of the Decisions Under Uncertainty with Allais' Rejoinder*, volume 21. Springer Science & Business Media, 2013.

S. Andersen, G. Harrison, M. Lau, and E. Rutstrom. Eliciting risk and time preferences. *Econometrica*, 76(3):583–618, 2008. ISSN 1468-0262.

J. Andreoni and C. Sprenger. Risk preferences are not time preferences. *American Economic Review*, 102(7):3357–76, 2012. doi: 10.1257/aer.102.7.3357. URL http://www.aeaweb.org/articles.php?doi=10.1257/aer.102.7.3357.

W. Arulampalam, A. L. Booth, and M. L. Bryan. Is there a glass ceiling over europe? exploring the gender pay gap across the wage distribution. *Industrial & Labor Relations Review*, 60(2):163–186, 2007.

N. Augenblick, M. Niederle, and C. Sprenger. Working over time: Dynamic inconsistency in real effort tasks. *Quarterly Journal of Economics*, forthcoming.

R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481 – 1509, 2004.

L. Barseghyan, J. Prince, and J. C. Teitelbaum. Are risk preferences stable across contexts? evidence from insurance data. *American Economic Review*, 101(2):591–631, 2011.

J. Beattie and G. Loomes. The impact of incentives upon risky choice experiments. *Journal of Risk and Uncertainty*, 14(2):155–68, March 1997. URL http://ideas.repec.org/a/kap/jrisku/v14y1997i2p155-68.html.

L. M. Berg, A.-S. Källberg, K. E. Göransson, J. Östergren, J. Florin, and A. Ehrenberg. Interruptions in emergency department work: an observational and interview study. *BMJ quality & safety*, 22(8):656–663, 2013.

M. H. Birnbaum and R. J. Gutierrez. Testing for intransitivity of preferences predicted by a lexicographic semi-order. *Organizational Behavior and Human Decision Processes*, 104 (1):96–112, 2007.

M. H. Birnbaum and U. Schmidt. Testing transitivity in choice under risk. *Theory and Decision*, 69(4):599–614, 2010.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience theory of choice under risk. Technical report, National Bureau of Economic Research, 2010.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience in experimental tests of the endowment effect. *American Economic Review*, 102(3):47–52, 2012.

C. Bradler and S. Neckermann. The magic of the personal touch: field experimental evidence on money and gratitude as gifts. *Working Paper*, 2015.

C. Bradler, R. Dur, S. Neckermann, and A. Non. Employee recognition and performance: A field experiment. *ZEW-Centre for European Economic Research Discussion Paper*, (13-017), 2013.

C. Bradler, S. Neckermann, and A. J. Warnke. Creativity is different: Comparing rewards across a creative and a routine task. *Working Paper*, 2015.

A. Brandon, J. List, S. Andersen, and U. Gneezy. Toward and understanding of reference-dependent labor supply: Theory and evidence from a field experiment. Technical report, The Field Experiments Website, 2014.

E. G. Brown. The hidden costs of interruptions at work. `http://www.fastcompany.com/3044667/work-smart/the-hidden-costs-of-interruptions-at-work`, 2015. Accessed: 2015-09-30.

J. Brown, T. Hossain, and J. Morgan. Shrouded attributes and information suppression: Evidence from the field. *Quarterly Journal of Economics*, 125(2), 2010.

M. Browning, A. Deaton, and M. Irish. A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica: journal of the econometric society*, pages 503–543, 1985.

M. S. Burke, J. R. Carter, R. D. Gominiak, and D. F. Ohl. An experimental note on the allais paradox and monetary incentives. *Empirical Economics*, 21(4):617–632, 1996.

T. Buser and N. Peter. Multitasking. *Experimental Economics*, 15(4):641–655, 2012.

C. Camerer, L. Babcock, G. Loewenstein, and R. Thaler. Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, pages 407–441, 1997.

C. F. Camerer. Does the basketball market believe in the 'hot hand'? *American Economic Review*, 79(5):pp. 1257–1261, 1989. ISSN 00028282. URL `http://www.jstor.org/stable/1831452`.

D. Card. Intertemporal Labor Supply: An Assessment. NBER Working Papers 3602, National Bureau of Economic Research, Inc, Jan. 1991. URL `http://ideas.repec.org/p/nbr/nberwo/3602.html`.

D. Card and D. Sullivan. Measuring the effect of subsidized training programs on movements in and out of employment. *Econometrica*, 56(3):497–530, 1988.

E. E. Cardella and B. Depew. Testing for the ratchet effect: Evidence from a real-effort work task. *Working Paper*, 2015.

S. Cerreia-Vioglio, D. Dillenberger, and P. Ortoleva. Cautious expected utility and the certainty effect. *Econometrica*, 83(2):693–728, 2015.

J. G. Chapela. On the price of recreation goods as a determinant of male labor supply. *Journal of Labor Economics*, 25(4):pp. 795–824, 2007. ISSN 0734306X. URL `http://www.jstor.org/stable/10.1086/519538`.

G. Charness. Attribution and reciprocity in an experimental labor market. *Journal of Labor Economics*, 22(3):665–688, 2004.

G. Charness and P. Kuhn. Lab labor: What can labor economists learn from the lab? *Handbook of labor economics*, 4:229–330, 2011.

G. Charness, D. Masclet, and M. C. Villeval. Competitive preferences and status as an incentive: Experimental evidence. *Groupe d'Analyse et de Theorie Economique Working Paper*, (1016), 2010.

X. Chen, J. Favilukis, and L. S. An estimation of economic models with recursive preferences. *Quantitative Economics*, 4:39–83, 2013.

H. W. Chesson and W. K. Viscusi. Commonalities in time and ambiguity aversion for long-term risks*. *Theory and Decision*, 54(1):57–71, 2003.

R. Chetty, A. Looney, and K. Kroft. Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177, 2009.

R. Chetty, A. Guren, D. Manoli, and A. Weber. Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *The American Economic Review*, 101(3):471–475, 2011.

M. Conlin, T. O'Donoghue, and T. J. Vogelsang. Projection bias in catalog orders. *The American Economic Review*, pages 1217–1249, 2007.

J. Conlisk. Three variants on the allais example. *American Economic Review*, 79(3):392–407, 1989.

M. Connolly. Here comes the rain again: Weather and the intertemporal substitution of leisure. *Journal of Labor Economics*, 26(1):pp. 73–100, 2008. ISSN 0734306X. URL http://www.jstor.org/stable/10.1086/522067.

B. Corgnet, R. Hernan-Gonzalez, and E. Schniter. Why real leisure really matters: incentive effects on real effort in the laboratory. *Experimental Economics*, pages 1–18, 2014. ISSN 1386-4157. doi: 10.1007/s10683-014-9401-4. URL http://dx.doi.org/10.1007/s10683-014-9401-4.

D. Coviello, A. Ichino, and N. Persico. Time allocation and task juggling. *The American Economic Review*, 104(2):609–623, 2014.

V. P. Crawford and J. Meng. New york city cab drivers' labor supply revisited: Reference-dependent preferences with rationalexpectations targets for hours and income. *The American Economic Review*, 101(5):1912–1932, 2011.

R. Cubitt, C. Starmer, and R. Sugden. On the validity of the random lottery incentive system. *Experimental Economics*, 1(2):115–131, September 1998. URL http://ideas.repec.org/a/kap/expeco/v1y1998i2p115-131.html.

M. Dean and P. Ortoleva. Is it all connected? a testing ground for unified theories of behavioral economics phenomena. mimeo, Columbia University, 2015.

R. Dhar and I. Simonson. The effect of the focus of comparison on consumer preferences. *Journal of Marketing Research*, pages 430–440, 1992.

D. L. Dickinson. An experimental examination of labor supply and work intensities. *Journal of Labor Economics*, 17(4):638–670, 1999.

L. Einav, A. Finkelstein, I. Pascu, and M. R. Cullen. How general are risk preferences? choices under uncertainty in different domains. *The American Economic Review*, 102(6): 2606–2638, 2012.

K. Eliaz and P. Ortoleva. Multidimensional ellsberg. *Management Science*, forthcoming.

F. Englmaier and S. Leider. Gift exchange in the lab-it is not (only) how much you give... 2010.

F. Englmaier and S. Leider. Managerial payoff and gift exchange in the field. *Available at SSRN 2175776*, 2012.

T. Epper and H. Fehr-Duda. The missing link: Unifying risk taking and time discounting. mimeo, University of Zurich, 2015.

T. Epper, H. Fehr-Duda, and A. Bruhin. Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. *Journal of Risk and Uncertainty*, 43(3):169–203, December 2011. URL `http://ideas.repec.org/a/kap/jrisku/v43y2011i3p169-203.html`.

L. Epstein and S. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57:937–969, 1989.

L. G. Epstein, E. Farhi, and T. Strzalecki. How much would you pay to resolve long-run risk? *American Economic Review*, 104(9):2680–97, 2014.

T. Eriksson, A. Poulsen, and M. C. Villeval. Feedback and incentives: Experimental evidence. *Labour Economics*, 16(6):679–688, 2009.

A. Falk and J. J. Heckman. Lab experiments are a major source of knowledge in the social sciences. *science*, 326(5952):535–538, 2009.

C.-P. Fan. Allais paradox in the small. *Journal of Economic Behavior & Organization*, 49 (3):411–421, 2002.

H. S. Farber. Is tomorrow another day? the labor supply of new york city cabdrivers. *Journal of Political Economy*, 113(1), 2005.

H. S. Farber. Reference-dependent preferences and labor supply: The case of new york city taxi drivers. *The American Economic Review*, 98(3):1069–1082, 2008.

E. Fehr and L. Goette. Do workers work more if wages are high? evidence from a randomized field experiment. *The American Economic Review*, 97(1):pp. 298–317, 2007. ISSN 00028282. URL `http://www.jstor.org/stable/30034396`.

E. Fehr and K. M. Schmidt. The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1:615–691, 2006.

E. Fehr, L. Goette, and C. Zehnder. A behavioral account of the labor market: The role of fairness concerns. *Institute for Empirical Research in Economics, University of Zurich, Working Paper*, (394), 2008.

D. Freeman, Y. Halevy, and T. Kneeland. Eliciting risk preferences using choice lists. mimeo, University of British Columbia, 2015.

X. Gabaix. A sparsity-based model of bounded rationality*. *Quarterly Journal of Economics*, 129(4), 2014.

X. Gabaix, D. Laibson, et al. Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics*, 121(2):505–540, 2006.

P. Ghirardato and M. Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research*, pages 864–890, 2001. ISSN 0364-765X.

D. Gill and V. Prowse. A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review*, 102(1):469–503, 2012.

D. Gill and V. L. Prowse. A novel computerized real effort task based on sliders. 2011.

U. Gneezy and J. A. List. Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica*, 74(5):1365–1384, 2006. ISSN 1468-0262. doi: 10.1111/j.1468-0262.2006.00707.x. URL `http://dx.doi.org/10.1111/j.1468-0262.2006.00707.x`.

C. Goldin. A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4):1091–1119, 2014.

V. M. Gonzalez and G. Mark. Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 113–120. ACM, 2004.

S. Grant, A. Kajii, and B. Polak. Many good choice axioms: when can many-good lotteries be treated as money lotteries? *Journal of Economic Theory*, 56(2):313–337, 1992.

F. Gul. A theory of disappointment aversion. *Econometrica*, 59(3):667–686, 1991.

Y. Halevy. Strotz meets Allais: Diminishing impatience and the certainty effect. *American Economic Review*, 98(3):1145–1162, 2008.

J. Hey and J. Lee. Do subjects separate (or are they sophisticated)? *Experimental Economics*, 8(3):233–265, September 2005. URL `http://ideas.repec.org/a/kap/expeco/v8y2005i3p233-265.html`.

C. Holt and S. Laury. Risk aversion and incentive effects. *American Economic Review*, 92 (5):1644–1655, 2002. ISSN 0002-8282.

C. A. Holt. Preference reversals and the independence axiom. *American Economic Review*, 76(3):508–15, June 1986. URL `http://ideas.repec.org/a/aea/aecrev/v76y1986i3p508-15.html`.

S. Huck and W. Müller. Allais for all: Revisiting the paradox in a large representative sample. *Journal of Risk and Uncertainty*, 44(3):261–293, 2012.

D. Huffman and L. Goette. Incentives and the allocation of effort over time: the joint role of affective and cognitive decision making. 2006.

D. Kahneman and A. Tversky. Prospect theory: an analysis of choice under risk. *Econometrica*, 47:263–291, 1979.

D. Kahneman, J. L. Knetsch, and R. H. Thaler. Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives*, 5(1):193–206, 1991.

E. Karni and Z. Safra. "preference reversal" and the observability of preferences by experimental methods. *Econometrica*, 55(3):675–685, 1987.

S. Kaur, M. Kremer, and S. Mullainathan. Self-control and the development of work arrangements. *The American Economic Review*, pages 624–628, 2010.

J. B. Kessler. When will there be gift exchange? addressing the lab-field debate with laboratory gift exchange experiments. 2013.

J. B. Kessler and M. I. Norton. Tax aversion in labor supply. *Journal of Economic Behavior & Organization*, 2015.

S.-H. Kim and W. Whitt. The power of alternative kolmogorov-smirnov tests based on transformations of the data. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 25(4):24, 2015.

J. L. Knetsch. The endowment effect and evidence of nonreversible indifference curves. *The American Economic Review*, 79(5):1277–1284, 1989.

M. Kosfeld and S. Neckermann. Getting more work for nothing? symbolic awards and worker performance. *American Economic Journal: Microeconomics*, pages 86–99, 2011.

B. Kőszegi and M. Rabin. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, pages 1133–1165, 2006.

B. Kőszegi and M. Rabin. Reference-dependent risk attitudes. *The American Economic Review*, pages 1047–1073, 2007.

B. Kőszegi and M. Rabin. Reference-dependent consumption plans. *The American Economic Review*, 99(3):909–936, 2009.

B. Kőszegi and A. Szeidl. A model of focusing in economic choice. *The Quarterly Journal of Economics*, 128(1):53–104, 2013.

M. Kräkel. Emotions in tournaments. *Journal of Economic Behavior & Organization*, 67 (1):204–214, 2008.

D. Kreps and E. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46(1):185–200, 1978.

S. Kube, M. A. Maréchal, and C. Puppe. The currency of reciprocity: Gift exchange in the workplace. *The American Economic Review*, pages 1644–1662, 2012.

H. Kurata, H. Izawa, and M. Okamura. Non-expected utility maximizers behave as if expected utility maximizers: An experimental test. *Journal of Socio-Economics*, 38(4): 622 – 629, 2009. ISSN 1053-5357. doi: 10.1016/j.socec.2008.12.010. URL `http://www.sciencedirect.com/science/article/pii/S1053535708002047`.

R. M. Landers, J. B. Rebitzer, and L. J. Taylor. Rat race redux: Adverse selection in the determination of work hours in law firms. *The American Economic Review*, pages 329–348, 1996.

S. D. Levitt and J. A. List. Viewpoint: On the generalizability of lab behaviour to the field. *Canadian Journal of Economics/Revue canadienne d'économique*, 40(2):347–370, 2007.

S. D. Levitt and S. Neckermann. What field experiments have and have not taught us about managing workers. *Oxford Review of Economic Policy*, 30(4):639–657, 2014. doi: 10.1093/oxrep/grv003. URL `http://oxrep.oxfordjournals.org/content/30/4/639.abstract`.

G. Loewenstein, T. O'Donoghue, and M. Rabin. Projection bias in predicting future utility. *The Quarterly Journal of Economics*, pages 1209–1248, 2003.

G. Loomes and R. Sugden. Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368):805–824, 1982.

F. A. Lozano. The Flexibility Of The Workweek In The United States: Evidence From The Fifa World Cup. *Economic Inquiry*, 49(2):512–529, 04 2011. URL `http://ideas.repec.org/a/bla/ecinqu/v49y2011i2p512-529.html`.

G. Mark, V. M. Gonzalez, and J. Harris. No task left behind?: examining the nature of fragmented work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 321–330. ACM, 2005.

G. Mark, D. Gudith, and U. Klocke. The cost of interrupted work: more speed and stress. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 107–110. ACM, 2008.

D. M. Mason and J. H. Schuenemeyer. A modified kolmogorov-smirnov test sensitive to tail alternatives. *The Annals of Statistics*, pages 933–946, 1983.

K. L. Milkman, J. A. Minson, and K. G. Volpp. Holding the hunger games hostage at the gym: An evaluation of temptation bundling. *Management Science*, 60(2):283–299, 2013.

J. Nakamura and M. Csikszentmihalyi. The concept of flow. *Handbook of positive psychology*, pages 89–105, 2002.

S. Nickell. Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426, 1981.

J. Nielsen. Im, not ip (information pollution). *Queue*, 1(8):76–75, 2003.

G. S. Oettinger. An empirical analysis of the daily labor supply of stadium venors. *Journal of political Economy*, 107(2):360–392, 1999.

S. Onay and A. Öncüler. Intertemporal choice under timing risk: An experimental approach. *Journal of Risk and Uncertainty*, 34(2):99–121, 2007.

P. Ortoleva. The price of flexibility: Towards a theory of thinking aversion. *Journal of Economic Theory*, 148(3):903–934, 2013.

L. Pistaferri. Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics*, 21(3):729–754, 2003.

D. G. Pope and M. E. Schweitzer. Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *The American Economic Review*, 101(1): 129–157, 2011.

D. G. Pope, J. Price, and J. Wolfers. Awareness reduces racial bias. Technical report, National Bureau of Economic Research, 2013.

J. Price and J. Wolfers. Racial discrimination among nba referees. *The Quarterly Journal of Economics*, 125(4):1859–1887, 2010. doi: 10.1162/qjec.2010.125.4.1859. URL `http://qje.oxfordjournals.org/content/125/4/1859.abstract`.

J. Quiggin. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, 1982.

M. Rabin. Incorporating fairness into game theory and economics. *The American economic review*, pages 1281–1302, 1993.

M. Rabin. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68:1281–1292, 2000.

Y. W. Ramírez and D. A. Nembhard. Measuring knowledge worker productivity: A taxonomy. *Journal of intellectual capital*, 5(4):602–628, 2004.

M. Regenwetter, J. Dana, and C. P. Davis-Stober. Transitivity of preferences. *Psychological Review*, 118(1):42, 2011.

A. Rubinstein. Similarity and decision-making under risk (is there a utility theory resolution to the allais paradox?). *Journal of Economic Theory*, 46(1):145–153, 1988.

O. Schaffer. Crafting fun user experiences: A method to facilitate flow. *Human Factors International. Online White paper. http://www. humanfactors. com/FunExperiences. asp*, 2013.

U. Segal. Two-stage lotteries without the reduction axiom. *Econometrica*, 58:349–377, 1990.

U. Segal and A. Spivak. First order versus second order risk aversion. *Journal of Economic Theory*, 51:111–125, 1990.

U. Simonsohn. Weather to go to college. *The Economic Journal*, 120(543):270–280, 2010.

J. B. Spira and J. B. Feintuch. The cost of not paying attention: how interruptions impact knowledge worker productivity. Technical report, September 2005.

R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, pages 165–180, 1955.

A. Tversky. Intransitivity of preferences. *Psychological review*, 76(1):31, 1969.

A. Tversky and D. Kahneman. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

M. R. Ward. Does time spent playing video games crowd out time spent studying? 2012.

M. Weigl, A. Müller, P. Angerer, and F. Hoffmann. Workflow interruptions and mental workload in hospital pediatricians: an observational study. *BMC health services research*, 14(1):433, 2014.

J. I. Westbrook, E. Coiera, W. T. Dunsmuir, B. M. Brown, N. Kelk, R. Paoloni, and C. Tran. The impact of interruptions on clinical task completion. *Quality and Safety in Health Care*, 19(4):284–289, 2010.

M. E. Yaari. The dual theory of choice under risk. *Econometrica*, 55:95–115, 1987.